

Université de Sherbrooke

**Séquençage d'un isolat canadien du virus de l'hépatite *delta*
humaine et développement d'outils pour l'étude de ce pathogène.**

Par
Mme Véronique Poisson
Département de Biochimie

Mémoire présenté à la Faculté de Médecine
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.) en Biochimie

2 décembre 1999



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-67709-5

À mon père...

TABLE DES MATIÈRES

<u>DÉDICACE</u>	i
<u>LISTE DES FIGURES</u>	v
<u>LISTE DES TABLEAUX</u>	vi
<u>LISTE DES ABRÉVIATIONS</u>	vii
<u>RÉSUMÉ</u>	ix
1.0 <u>INTRODUCTION</u>	1
1.1 HÉPATITES INFECTIEUSES	1
1.1.1 Le virus de l'hépatite A	2
1.1.2 Le virus de l'hépatite B	3
1.1.3 Le virus de l'hépatite C	6
1.1.4 Le virus de l'hépatite E	7
1.1.5 Le virus de l'hépatite G	9
1.2 VIRUS DE L'HÉPATITE <i>DELTA</i> HUMAINE	9
1.2.1 Historique	10
1.2.2 Hépatite infectieuse de type D	11
1.2.2.1 Co-infection VHB/VHD	12
1.2.2.2 Surinfection	14
1.2.3 Épidémiologie	17
1.3 CARACTÉRISTIQUES DU VIRUS DE L'HÉPATITE <i>DELTA</i> HUMAINE	19
1.3.1 Structure du virion	19
1.3.2 Le génome viral et les génotypes	21
1.3.3 Pathogenèse	24
1.3.4 Mécanisme de réplication virale	25
1.3.4.1 Les domaines autocatalytiques (ribozymes <i>delta</i>)	27
1.3.4.2 Les antigènes <i>delta</i>	29
1.3.5 Taxonomie	29
1.4 APERÇU ET OBJECTIFS DE LA RECHERCHE	32
2.0 <u>MATÉRIEL ET MÉTHODES</u>	35
2.1 ALIGNEMENT ET ANALYSE DES SÉQUENCES COMPLÈTES DU VHD	35

2.2 SÉRUM D'UN PORTEUR CANADIEN	36
2.3 ARN VIRAL DU VIRUS DE L'HÉPATITE <i>DELTA</i> HUMAINE	36
2.3.1 Extraction de l'ARN viral	36
2.3.2 Détermination et synthèse des amorces RT-PCR	37
2.3.3 Synthèse de l'ADNc	38
2.3.4 Amplification de différents fragments d'ADNc	39
2.3.5 Purification des différents fragments d'ADNc	39
2.3.6 Clonage des différents fragments d'ADNc	40
2.3.7 Séquençage des différents fragments d'ADNc	41
2.4 DÉVELOPPEMENT D'UNE LIGNÉE CELLULAIRE HepG2	42
2.4.1 Mise en culture des cellules	42
2.4.2 Transfection stable des vecteurs d'expression pcDNA ₃ et pcDNA ₃ -Ag-S et isolation des colonies cellulaires	42
2.5 CARACTÉRISATION DES COLONIES EXPRIMANT L'ARN AgHD-S	43
2.5.1 Extraction d'ARN cytoplasmique	43
2.5.2 Synthèse des ADNc respectifs des ARN AgHD-S et GAPDH	44
2.5.3 Amplification des ADNc AgHD-S et GAPDH	45
2.5.4 Buvarpage de type Southern	46
2.5.5 Synthèse <i>in vitro</i> et purification de la sonde d'ARN HDAG-S	47
2.5.6 Synthèse <i>in vitro</i> de la sonde d'ADN GAPDH	49
2.5.7 Hybridation des sondes	50
3.0 <u>RÉSULTATS</u>	51
3.1 COMPILATION DE SÉQUENCES	51
3.2 ALIGNEMENT ET ANALYSE DES SÉQUENCES COMPLÈTES DU VHD	54
3.3 SÉQUENÇAGE D'UN ISOLAT CANADIEN DU VHD	58
3.3.1 Stratégie de clonage	58
3.3.2 Clonage et séquençage des produits d'amplification RT-PCR	63
3.3.3 Analyse de la séquence virale de l'isolat canadien du VHD	65
3.4 CARACTÉRISATION ET IDENTIFICATION DES COLONIES CELLULAIRES EXPRIMANT CONSTITUTIVEMENT L'ARN AgHD-S	66

4.0 <u>DISCUSSION</u>	71
4.1 SÉQUENCE DU VHD	72
4.2 LIGNÉE CELLULAIRE	74
5.0 <u>REMERCIEMENTS</u>	78
6.0 <u>RÉFÉRENCES</u>	79
7.0 <u>ANNEXES</u>	84
7.1 ÉLECTROPHORÈSE EN GEL D'URÉE-POLYACRYLAMIDE	85
7.2 ALIGNEMENT DES 23 SÉQUENCES COMPLÈTES DU VHD	86
7.3 ARTICLE 1: A canadian isolate of hepatitis <i>delta</i> virus	101
7.4 ARTICLE 2: The viroid and viroid-like RNA database (1999)	118
7.5 ARTICLE 3: The viroid and viroid-like RNA database (1998)	121

LISTE DES FIGURES

Figure 1. Marqueurs sérologiques d'une co-infection VHB/VHD	13
Figure 2. Marqueurs sérologiques d'une surinfection par le VHD d'un porteur chronique du VHB	15
Figure 3. Structure de la particule virale <i>delta</i>	20
Figure 4. Représentation schématique de la structure secondaire en tige du génome du virus de l'hépatite <i>delta</i>	23
Figure 5. Représentation du cycle de réplication en cercle roulant du VHD	26
Figure 6. Représentation schématique des motifs autocatalytiques d'ARN du VHD	28
Figure 7. Compilation de séquences et informations relatives au VHD	53
Figure 8. Représentation schématique des trois domaines retrouvés au sein du génome viral du VHD	57
Figure 9. Représentation schématique des fragments amplifiés par PCR	60
Figure 10. Fragments d'amplification obtenus par RT-PCR	62
Figure 11. Séquence d'un isolat canadien du VHD	64
Figure 12. RT-PCR réalisés à partir de l'ARN cytoplasmique extrait des colonies cellulaires	70

LISTE DES TABLEAUX

Tableau 1. Similitudes et différences entre le VHD et les viroïdes	31
Tableau 2. Séquence et position des amorces de RT-PCR et de séquençage	59

LISTE DES ABRÉVIATIONS

A	adénine ou (désoxy) riboadénosine mono-, di- ou triphosphate
a.a.	acide aminé
ADN	acide désoxyribonucléique
ADNc	acide désoxyribonucléique complémentaire
AgHD	antigène du virus de l'hépatite <i>delta</i>
AgHD-S	antigène du virus de l'hépatite <i>delta</i> , forme petite
AgHD-L	antigène du virus de l'hépatite <i>delta</i> , forme large
ARN	acide ribonucléique
ARNm	acide ribonucléique messenger
bp	bleu de bromophénol
C	cytosine ou (désoxy) ribocytidine mono-, di- ou triphosphate
cpm	compte par minute
db	double-brin d'ADN ou d'ARN
DEPC	diéthyl pyrocarbonate
DTT	dithiothréitol
dNTP	désoxyribonucléotide triphosphate ou N= A, C, G, T
EDTA	acide éthylènediaminetétraacétique
G	guanine ou (désoxy) riboguanosine mono-, di- ou triphosphate
G-418	Geneticin™
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase d'ADN ou d'ARN
kDa	kilodalton
<i>lacZ</i>	gène de la β-galactosidase
MEM	<i>Eagle's Minimal Essential Medium</i>
nt	nucléotide
NTP	ribonucléotide triphosphate ou N= A, C, G, U
ORF	cadre de lecture ouvert
pb	paire de base d'ADN ou d'ARN
PCR	réaction de polymérisation en chaîne
PSA	persulfate d'ammonium
RT	transcription inverse
RT-PCR	transcription inverse couplée à une réaction de polymérisation en chaîne
sb	simple-brin d'ADN ou d'ARN
SDS	sodium dodécyl sulfate
SFB	sérum fœtal bovin
SV40	<i>simian virus 40</i>

T	thymine ou (désoxy) thymidine mono-, di- ou triphosphate
TAE	Tris-Acétate-EDTA
TBE	Tris-Borate-EDTA
TEMED	N, N, N' N'-tétra-méthyléthylènediamine
TRIS	hydroxyméthyl-aminométhane
U	uracile ou uridine mono-, di-, ou triphosphate
UV	ultraviolet
VHA	virus de l'hépatite A humaine
VHB	virus de l'hépatite B humaine
VHC	virus de l'hépatite C humaine
VHD	virus de l'hépatite D humaine
VHE	virus de l'hépatite E humaine
VHG	virus de l'hépatite G humaine
VIH	virus d'immunodéficience humaine
v/v	volume/volume
xc	xylène cyanol
Xgal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

RÉSUMÉ

Séquençage d'un isolat canadien du virus de l'hépatite *delta* humaine et développement d'outils pour l'étude de ce pathogène.

Par Mme Véronique Poisson
Département de Biochimie
Université de Sherbrooke

Mémoire présenté à la Faculté de Médecine
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.) en Biochimie

2 décembre 1999

Le virus de l'hépatite D (VHD), autrefois appelé agent *delta*, infecte plus de 15 millions d'individus à travers le monde. Cet agent infectieux requiert la présence du virus de l'hépatite B (VHB) pour former son enveloppe. L'infection par le VHD cause des hépatites fulminantes et chroniques avec cirrhose du foie. Le VHD est un petit virus de forme circulaire dont le diamètre est de 36 nm. Le génome viral est constitué d'un ARN circulaire de 1,7 kilobases qui présente beaucoup de complémentarité interne le conduisant à se replier en une structure semblable à une tige. Il existe beaucoup de variations de séquence du génome du VHD et plusieurs génotypes différents ont été identifiés. Chez un patient infecté par un même génotype, la population virale est aussi hétérogène (quasi-espèce). Possiblement en raison de cette diversité génétique, le VHD réussit à déjouer la surveillance immunitaire de l'hôte, démontre des symptômes variés, échappe souvent au traitement et rend le développement d'un vaccin difficile. Le VHD possède plusieurs caractéristiques uniques qui en font un sujet d'étude très intéressant. L'objectif de ce projet de recherche est de développer des outils pour faciliter l'étude et la compréhension des propriétés biologiques de ce virus. D'abord, pour contribuer à l'étude des mécanismes d'évolution et approfondir nos connaissances sur l'aspect moléculaire du VHD, nous avons procédé à l'alignement précis des séquences nucléotidiques complètes connues de ce virus ainsi qu'à l'analyse de cet alignement. Les séquences et l'alignement furent mis à la disposition de la communauté scientifique sous forme d'une banque de données. Celle-ci constitue un outil, entre autre, pour l'étude des concepts d'hétérogénéité, de microhétérogénéité et de quasi-espèce, relatifs au VHD et autres ARNs apparentés. En deuxième lieu, l'établissement d'une relation entre les génotypes, le degré de variabilité de séquences, la distribution géographique et le niveau de la sévérité de la maladie nécessite un grand nombre de séquences virales du VHD. C'est pourquoi nous avons déterminé la séquence nucléotidique complète du VHD isolé d'un individu canadien. L'ADNc de l'ARN viral extrait d'un sérum fut amplifié par PCR, cloné et séquencé. De façon à optimiser la synthèse de l'ADNc et l'amplification subséquente, nous avons déterminé des amorces correspondant à des régions très conservées du génome viral à l'aide de l'alignement de séquences. Une seule amorce fut nécessaire pour

effectuer la synthèse de l'ADNc. L'analyse de la séquence a révélé que ce nouvel isolat contenant 1672 nucléotides correspond au génotype I. Le séquençage de quatre clones indépendants a permis de déterminer 17 substitutions, correspondant à une variabilité de séquence moyenne de 1%. Sept de ces mutations furent retrouvées dans la petite région constituée de 48 nucléotides localisée entre les deux motifs autocatalytiques hautement conservés permettant de la définir comme une région hypervariable. Cette nouvelle séquence du VHD a été ajoutée à l'alignement et à la banque de données que possède le laboratoire. Enfin, pour faciliter l'étude des propriétés biochimiques de ce pathogène, une lignée cellulaire HepG2 fut transfectée de façon stable par un vecteur permettant l'expression constitutive de l'ARNm correspondant à la forme petite de l'antigène *delta* (AgHD-S). L'expression de l'ARNm fut détectée par RT-PCR et confirmée par buvardage de type Southern. Ces analyses ont permis l'identification de deux colonies cellulaires exprimant constitutivement l'ARNm de l'AgHD-S. En bref, ce travail a permis d'approfondir nos connaissances sur l'aspect moléculaire du VHD et le développement d'outils pouvant être utiles pour les travaux visant, entre autre, l'inhibition de la propagation de ce pathogène

1.0 INTRODUCTION

Depuis toujours, l'homme n'a cessé d'être infecté par des virus. Ces derniers sont des agents infectieux dont la structure se résume à un génome protégé par une coque protéique (nucléocapside) qui est entourée d'une enveloppe chez certains virus. En raison de leur extrême simplicité, les virus sont des parasites intracellulaires absolus des cellules eucaryotes. Malgré leur simplicité, ils continuent chaque jour de faire de nombreuses victimes. Les virus ne peuvent donc être ignorés par la communauté scientifique.

1.1 HÉPATITES INFECTIEUSES

Par leur fréquence, combinée à leur gravité évolutive potentielle, les hépatites virales constituent une des toutes premières maladies infectieuses d'origine virale. Le terme "hépatite virale" est utilisé pour décrire l'atteinte hépatique causée par un nombre limité de virus hépatotropes, c'est-à-dire infectieux pour le foie et dont la manifestation clinique principale est une hépatite. À ce jour, les hépatites virales sont dues essentiellement à 6 virus (A, B, C, D, E et G), qui diffèrent suivant leur mode de transmission et leur évolution chronique ou non. Notre connaissance des hépatites virales, en permanente évolution, a abouti depuis 20 ans à des progrès spectaculaires sur les plans diagnostiques, thérapeutique et prophylactique. Aujourd'hui, la recherche de nouveaux virus hépatotropes se poursuit, particulièrement en présence d'une hépatite fulminante, où environ 50% des cas ne peuvent être attribués à aucun des virus déjà connus (Trépo et Valla, 1993). Les virus identifiés à ce jour n'expliquent apparemment pas toutes les

hépatites aiguës ou chroniques et laissent encore la place à d'autres agents transmissibles responsables d'hépatites.

1.1.1 Le virus de l'hépatite A

Découvert en 1973, le virus de l'hépatite A (VHA) est un virus non enveloppé, de 27 nm de diamètre, d'aspect sphérique et à surface lisse appartenant à la famille des *Picornaviridae* (Janot, 1996). Le génome du VHA est un ARN linéaire simple-brin d'une longueur de 7,5 kb. Ce virus se transmet habituellement par voie fécale-orale directe, l'homme étant le seul réservoir. L'eau, les fruits de mer et certains aliments peuvent être impliqués indirectement et provoquer d'importantes épidémies. La plupart des cas d'infection par le VHA sont sporadiques ou surviennent par petites épidémies. La maladie est commune dans les milieux surpeuplés où l'hygiène individuelle et collective est souvent mal contrôlée. La plupart des cas d'hépatite A représentent la forme la plus bénigne de la maladie. Toutefois, des cas sévères sont possibles, surtout si l'individu est également porteur du virus de l'hépatite C. La période d'incubation varie de 2 à 7 semaines. L'hépatite A est généralement asymptomatique. Dans les autres cas, le tableau clinique peut être caractérisé par des malaises généraux, de la fièvre, une asthénie importante, une anorexie, une hépatalgie, des nausées, des vomissements, des urines foncées et l'apparition d'un ictère. L'infection est d'une durée variable mais n'évolue jamais vers la chronicité et la majorité des patients n'ont pas de séquelle.

Le marqueur le plus sensible et le plus spécifique d'une hépatite aiguë de type A est l'anticorps IgM anti-VHA. Le diagnostic direct par isolement du virus ou du matériel génétique viral est du domaine de la recherche. La pathogenèse du VHA de même que l'origine des lésions hépatiques ne sont toujours pas connues avec certitude, mais il apparaît de plus en plus que des mécanismes immunologiques soient en cause. Il n'existe pas de traitement spécifique, cependant une prévention efficace peut être réalisée par la mise en oeuvre de mesures d'hygiène appropriées. L'immunisation est recommandée pour tout voyageur qui prévoit séjourner dans les pays où l'endémicité est modérée ou élevée. Les immunoglobulines (immunisation passive) et plus récemment, le vaccin (immunisation active) sont disponibles.

1.1.2 Le virus de l'hépatite B

Le virus de l'hépatite B (VHB), identifié en 1963 par Blumberg, est un virus enveloppé de 42 nm appartenant à la famille des *Hepadnaviridae* (Trépo et Valla, 1993). Le nombre de porteurs chroniques asymptomatiques à travers la population mondiale est évalué à plus de 300 millions. Une enveloppe externe (AgHBs) entoure le noyau central (AgHBc) du virus. Ce noyau contient le génome viral constitué d'ADN circulaire et partiellement double-brin d'une longueur de 3,2 kb. Le brin le plus long comprend quatre cadres de lecture encodant respectivement les protéines de l'enveloppe pré-S et S (AgHBs), les protéines de la capside pré-C et C (AgHBe et AgHBc), la polymérase et une protéine X dont le rôle n'est pas parfaitement élucidé. Le VHB se réplique par des

mécanismes similaires à ceux des rétrovirus. L'ADN polymérase du VHB possède en effet une activité transcriptase inverse (ADN-ARN-ADN). Il est admis que le VHB a peu d'effet cytopathogène, les lésions hépatiques étant, en réalité, la conséquence d'un ensemble de réactions immunologiques à médiation principalement cellulaire. Ces réactions sont dirigées contre les hépatocytes dont la membrane exprime les antigènes de la nucléocapside virale, cette expression impliquant une réplication complète du VHB.

Les trois modes majeurs de transmission de l'hépatite B sont la transmission parentérale, sexuelle et périnatale. Par conséquent, le virus se retrouve dans le sang, le sperme, le flux menstruel et les autres liquides de l'organisme d'une personne infectée par le VHB. 5-10% des adultes et près de 90% des bébés qui souffrent d'une hépatite B vont demeurer porteurs du virus pour le restant de leur vie. Ces porteurs chroniques pourront transmettre le virus à d'autres personnes. Le tableau clinique est très variable, allant de la forme asymptomatique à l'hépatite aiguë cliniquement apparente jusqu'à la forme fulminante. La période d'incubation est en moyenne de 10 semaines. Le tableau clinique est caractérisé par un début graduel de fatigue, anorexie, nausées, arthralgies avec ou sans arthrite, décoloration des selles, urines foncées et jaunisse. Les symptômes de l'hépatite B sont souvent plus sévères que ceux présentés lors d'une hépatite A. La gravité de l'hépatite virale B est essentiellement liée à la possibilité d'une évolution chronique de l'infection virale où environ 10% des cas vont évoluer vers la forme chronique d'infection. Celle-ci peut être asymptomatique ou évoluer vers une hépatite chronique active. Au risque d'évolution vers la cirrhose qui est de l'ordre de 20% en 10 à 20 ans chez les patients ayant une hépatite chronique active, s'ajoute la possibilité de développer

un carcinome hépatocellulaire. Le risque de développer cette tumeur est de l'ordre de 3 à 5% par an chez les patients cirrhotiques.

Le sérodiagnostic d'une hépatite B aiguë est fondé sur l'utilisation de deux marqueurs, soit l'antigène HBs (AgHBs) et l'IgM anti-HBc. L'hépatite chronique B est reconnue habituellement par la persistance des signes cliniques, biologiques et l'AgHBs au-delà de six mois. La physiopathologie de l'hépatite B est mal connue mais il semble que les lésions hépatiques seraient le résultat de la réaction immunologique contre l'infection virale plutôt que le résultat d'un pouvoir cytopathogène direct du virus. Le traitement par certains agents antiviraux comme la lamivudine, qui bloque la transcriptase inverse, se sont avérés efficaces. De plus, l'interféron-alpha permet d'influer sur l'histoire naturelle de l'hépatite B chronique en raccourcissant sa durée. Donc, ce traitement permet d'éviter, dans un certain nombre de cas, l'évolution vers la cirrhose et probablement d'empêcher l'apparition du carcinome hépatocellulaire. Cependant, le traitement par l'interféron-alpha s'est avéré efficace chez moins de 40 p. 100 des patients, n'est disponible qu'en injections, est associé à plusieurs effets secondaires et est très dispendieux. L'utilisation de la lamivudine s'est avérée bien tolérée et efficace même chez des patients qui n'avaient pas répondu à l'interféron-alpha. Des études complémentaires dans lesquelles une association de lamivudine et de l'interféron-alpha est utilisée sont en cours. En terme de prévention, la prophylaxie passive et active de l'hépatite B est disponible.

1.1.3 Le virus de l'hépatite C

Depuis l'élimination des dons de sang provenant des sujets porteurs de l'AgHBs, les hépatites non A non B se sont avérées être la principale cause des hépatites posttransfusionnelles. Dès lors, des expériences ont montré qu'il s'agissait du virus de l'hépatite C (VHC). Découvert par le groupe de Houghton en collaboration avec l'équipe de Bradley (Choo et al., 1989), le VHC est un virus enveloppé d'environ 50 nm de diamètre. Ce virus possède un génome d'ARN simple-brin d'une longueur de 9,6 kb et fait partie de la famille des *Flaviviridae* (Janot, 1996). Il y a beaucoup de variations dans le génome du VHC et plusieurs génotypes différents ont été identifiés. Chez un patient infecté par un même génotype, la population virale est aussi hétérogène (quasi-espèce). Possiblement en raison de cette diversité génétique, le VHC réussit à déjouer la surveillance immunitaire de l'hôte, démontre un haut taux de chronicité, échappe souvent au traitement et rend le développement d'un vaccin difficile.

Les connaissances sur l'épidémiologie du VHC furent longtemps limitées en raison de l'échec de l'isolement d'un agent causal de la maladie. L'année 1989 représente une étape décisive avec l'identification du génome du VHC et le développement de tests sérologiques spécifiques autorisant le dépistage d'anticorps antiviraux correspondants. Le VHC se transmet essentiellement par voie sanguine et percutanée, il est à l'origine de plus de 85% des hépatites posttransfusionnelles. Les manifestations cliniques ressemblent à celles de l'hépatite B et la période d'incubation est d'environ 6 semaines. Les infections

aiguës à VHC sont très fréquemment asymptomatiques ou présentent à l'occasion des symptômes non spécifiques tels une anorexie légère, nausée, fièvre, frissons, malaises généraux, fatigue, hépatalgie et occasionnellement des selles décolorées et des urines foncées. L'ictère n'est présent que dans 25% des cas aigus. Les infections à VHC évoluent vers la chronicité dans plus de 50% des cas. L'infection chronique par le VHC est associée à un risque accru de carcinome hépatocellulaire surtout chez les patients cirrhotiques. La physiopathologie des hépatites C est mal connue. On ne sait pas quelle est la part respective de l'effet cytopathogène et de la réaction immunologique contre l'infection virale.

Le diagnostic sérologique de la maladie se confirme par la détection des anticorps (anti-VHC) dans le sérum du patient. Chez la plupart des patients avec hépatite C chronique, on peut aussi détecter l'ARN du VHC par RT-PCR. Un traitement réalisé à l'aide de l'interféron-alpha a montré son efficacité au cours des hépatites chroniques C. Par contre, les patients porteurs d'une cirrhose répondent moins bien au traitement par l'interféron-alpha. De plus, ce traitement est très dispendieux.

1.1.4 Le virus de l'hépatite E

Le virus de l'hépatite E (VHE) est un virus sphérique, non enveloppé qui appartient à la famille des *Caliciviridae* (Trépo et Valla, 1993). Le génome viral est composé d'un ARN simple-brin d'une longueur de 7,5 kb. Le VHE fut ainsi nommé lors de la découverte du virus responsable d'épidémies d'hépatites non A, non B transmises par voie

fécale-orale. L'épidémiologie de l'hépatite E ressemble beaucoup à celle de l'hépatite A puisque les deux sont transmises par voie fécale-orale et la maladie a tendance à survenir soit sous forme d'épidémies ou sous forme sporadique. Le VHE se caractérise par l'atteinte préférentielle de l'adolescent et du jeune adulte. Les épidémies par le VHE frappent les pays en voie de développement où l'hygiène de l'eau est aléatoire. Hormis les formes graves, le VHE se caractérise par une évolution vers la guérison en quelques semaines, sans séquelle et sans passage à la chronicité. L'hépatite E se distingue par le risque d'une forme fulminante plus élevé que pour les autres types d'hépatites virales. Ceci fut particulièrement observé chez les femmes enceintes. Le tableau clinique est superposable à celui de l'hépatite A. En attendant la diffusion de tests sérologiques spécifiques, son diagnostic repose sur la négativité des tests d'identification des autres virus hépatotropes. Néanmoins, des techniques immunologiques en microscopie électronique et des essais enzymatiques visant la mise en évidence des anticorps contre le virus de l'hépatite E (anti-VHE) et l'antigène du VHE sont à l'étude. L'immunisation passive par les immunoglobulines n'a pas fait preuve d'efficacité. Par contre, le foisonnement des travaux auxquels le séquençage du VHE a ouvert la voie laisse présager une évolution rapide des connaissances de la maladie et la mise au point d'un vaccin au cours des prochaines années.

1.1.5 Le virus de l'hépatite G

Le virus de l'hépatite G (VHG) est un virus enveloppé appartenant à la famille des *Flaviviridae* qui a été mis en évidence en 1995 par la firme Genelabs. D'autres virus ont été découverts par la firme Abpott. Il s'agirait seulement de souches différentes d'un même virus (Simons et al., 1995; Muerhoff et al., 1995). La transmission du VHG se fait essentiellement par voie parentérale. Il y aurait une prévalence élevée du VHG chez les patients transfusés et les utilisateurs de drogues injectables, souvent en association avec une infection par le VHC. La plupart des patients infectés par le VHG deviennent des porteurs chroniques. Cependant, la pathogénicité du VHG semble très faible; la persistance du VHG ne cause que rarement des lésions histologiques et n'aggrave pas l'évolution d'une infection concurrente par le VHC. En raison de la découverte récente du virus de l'hépatite G, les connaissances sur ce dernier sont encore limitées.

1.2 VIRUS DE L'HÉPATITE *DELTA* HUMAINE

Autrefois appelé agent *delta*, le VHD est responsable de l'une des 6 formes d'hépatite virale chez l'humain connues jusqu'à ce jour (revue par Hadziyannis, 1997; Carreno et al., 1994; Lai, 1995; Lazinski et Taylor, 1994). À travers le monde, plus de 15 millions d'individus sont porteurs du VHD. Malheureusement, aucun vaccin ou traitement n'existe pour combattre l'infection par le VHD, ce qui constitue un problème de santé publique important. La prévalence de ce virus est relativement élevée sur les

continents sud-américain, africain et asiatique mais est plutôt faible en Amérique du Nord et en Europe, mise à part le bassin méditerranéen. Le VHD se transmet principalement par voie parentérale et sa prévalence est élevée chez les hémophiles et les utilisateurs de drogues injectables. Le VHD est hépatotropique, c'est-à-dire qu'il infecte seulement les hépatocytes. L'infection par le VHD peut provoquer une hépatite aiguë pouvant être fulminante et augmente considérablement les risques d'hépatocarcinome (revue par Lazinski et Taylor, 1994).

1.2.1 Historique

En 1977 à Turin, le Dr Mario Rizzetto, un gastro-entérologue, et ses collègues firent la découverte d'un nouvel antigène nucléaire dans les hépatocytes provenant de biopsies hépatiques de patients infectés par le virus de l'hépatite B (revue par Lazinski et Taylor, 1994). Cet antigène, nommé alors *delta* (AgHD) avait des anticorps correspondants retrouvés dans le sérum des patients souffrant d'hépatites B chroniques. Dans un premier temps, on pensait que l'AgHD était un nouvel antigène de la capside du VHB. Quelques années plus tard, des études dans lesquelles on procédait à des transmissions expérimentales à des chimpanzés ont permis de comprendre que l'AgHD était dérivé d'un autre virus, le VHD, et nécessitait le VHB pour causer une infection. Bien que l'antigène *delta* est un agent distinct et ne présente pas de similarité avec les antigènes du VHB, il est associé au VHB puisque l'antigène *delta* n'est détecté que chez les patients qui ont le marqueur AgHBs. En fait, la particule virale du VHD est entourée

d'une enveloppe composée de l'antigène de surface (AgHBs) du VHB et une nucléocapside formée du génome viral et de l'antigène *delta* (AgHD), l'enveloppe virale étant nécessaire pour avoir une particule virale complète et infectieuse du VHD (revue par Lai, 1995).

1.2.2 Hépatite infectieuse de type D

Empruntant les protéines AgHBs du VHB pour former son enveloppe, le VHD ne peut infecter un sujet qu'en association avec celui-ci. Cette association peut se faire de deux façons. Des études de transmissions expérimentales du VHB et du VHD à des chimpanzés ont permis de mieux comprendre les deux types d'association VHB/VHD (revue par Lazinski et Taylor, 1994). L'infection par le VHD peut survenir par co-infection avec le VHB ou par surinfection d'un sujet chroniquement infecté par le VHB. Du fait de la présence de cellules hépatiques produisant de l'AgHBs chez un porteur chronique du VHB, le VHD se trouve dans un milieu approprié pour sa propagation et des quantités infimes de VHD sont suffisantes pour occasionner la surinfection. À la différence du VHB et des autres hepadnavirus qui infectent *in vivo* des tissus extrahépatiques comme le rein, la rate et les lymphocytes circulants, le VHD montre un tropisme exclusivement hépatique.

1.2.2.1 Co-infection VHB/VHD

Cliniquement, rien ne permet de distinguer une hépatite aiguë par co-infection VHB/VHD d'une hépatite aiguë par le VHB seul. L'association des deux virus augmente le risque d'hépatite fulminante. Paradoxalement, le pronostic de l'hépatite B-*delta* est meilleur que celui de l'hépatite B fulminante. La présence du VHD n'augmente pas le risque de développer une infection chronique par le VHB car l'arrêt de la réplication du VHB entraîne celle du VHD. L'infection par le VHD est donc limitée par celle du VHB. Le diagnostic de la co-infection repose sur l'apparition concomitante de l'antigène HBs et d'IgM anti-HBc de titre élevé (témoin d'une infection récente par le VHB) et des marqueurs du VHD. Toutefois, comme le VHD a la capacité d'inhiber la réplication du VHB, l'AgHBs peut s'avérer indétectable dans le sérum. Les marqueurs du VHD peuvent donc être recherchés dans le cas d'une hépatite aiguë AgHBs négative/IgM anti-HBc positive. Les premiers marqueurs du VHD détectables dans le sérum sont l'AgHD par une méthode immunoenzymatique et l'ARN viral par hybridation moléculaire ou par RT-PCR. Ces marqueurs sont présents pendant une à quatre semaines. Les anticorps anti-HD de classe IgM apparaissent transitoirement après deux à cinq semaines d'évolution. Ces IgM font ensuite place à des IgG qui persistent à faible titre (10^{-1} ou 10^{-2}) (Figure 1.). Dans les rares cas (moins de 5% des cas de co-infection) où une infection chronique du VHB s'installe, il s'ensuit une infection chronique par le VHD dont l'évolution sérologique ressemble à la surinfection. La chronicité du VHD s'explique par la présence de l'AgHBs en grande quantité constituant ainsi un milieu favorable à la propagation du VHD.

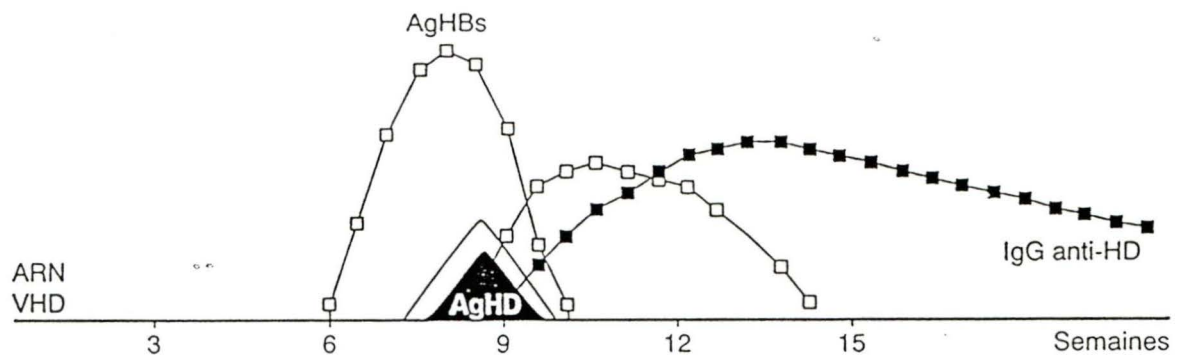


Figure 1: Marqueurs sérologiques d'une co-infection VHB/VHD. Après deux à six semaines d'infection, la présence d'AgHBs et d'IgM anti-HBc de titre élevé et des marqueurs du VHD indique une co-infection. Les marqueurs du VHD peuvent aussi être recherchés dans le cas d'une hépatite aiguë AgHBs négative/IgM anti-HBc positive. Les premiers marqueurs du VHD détectables dans le sérum sont l'AgHD et l'ARN présents pendant une à quatre semaines. Les anticorps anti-HD de classe IgM apparaissent après deux à cinq semaines d'évolution. Ces IgM font ensuite place à des IgG qui persistent à faible titre (10^{-1} ou 10^{-2}). (Trépo et Valla, 1993).

1.2.2.2 Surinfection

Un surinfection par le VHD d'un sujet chroniquement infecté par le VHB provoque une hépatite aiguë pouvant devenir fulminante. Dans plus de 80% des cas, il y a développement d'une infection chronique *delta* sévère. La figure 2 montre l'évolution sérologique classique d'une hépatite chronique *delta*. Si l'infection chronique par le VHB est connue, le diagnostic de la surinfection *delta* se caractérise par l'apparition de marqueurs du VHD. En absence de la connaissance d'une sérologie VHB antérieure, le diagnostic repose sur la mise en évidence des marqueurs du VHD et de l'AgHBs, en l'absence d'IgM anti-HBc, témoignant d'un contact ancien avec le VHB. Toutefois, le VHD inhibant la réplication du VHB, l'AgHBs peut être indétectable pendant la phase aiguë (début de l'infection) de la surinfection et parfois même pendant plusieurs mois. Cette éventualité doit être prise en considération dans la démarche diagnostique d'une hépatite virale AgHBs négative. Il existe une réponse anticorps anti-*delta* importante avec apparition d'IgM, puis d'IgG coexistant avec les marqueurs d'une hépatite B chronique.

Au cours de l'hépatite chronique *delta*, les titres d'IgM et d'IgG anti-*delta* demeurent généralement élevés et stables et persistent généralement plusieurs semaines, voir plusieurs mois dans le sérum. Il semble que les IgM anti-HD constituent un marqueur de l'agressivité de la maladie hépatique causée par le VHD. Les marqueurs complémentaires de l'infection chronique causée par le VHD sont la détection immunocytochimique de l'AgHD intrahépatique et la mise en évidence de l'ARN viral du

VHD dans le sérum. Au stade chronique *delta*, le diagnostic repose aussi sur la ponction-biopsie hépatique qui démontre souvent des lésions sévères.

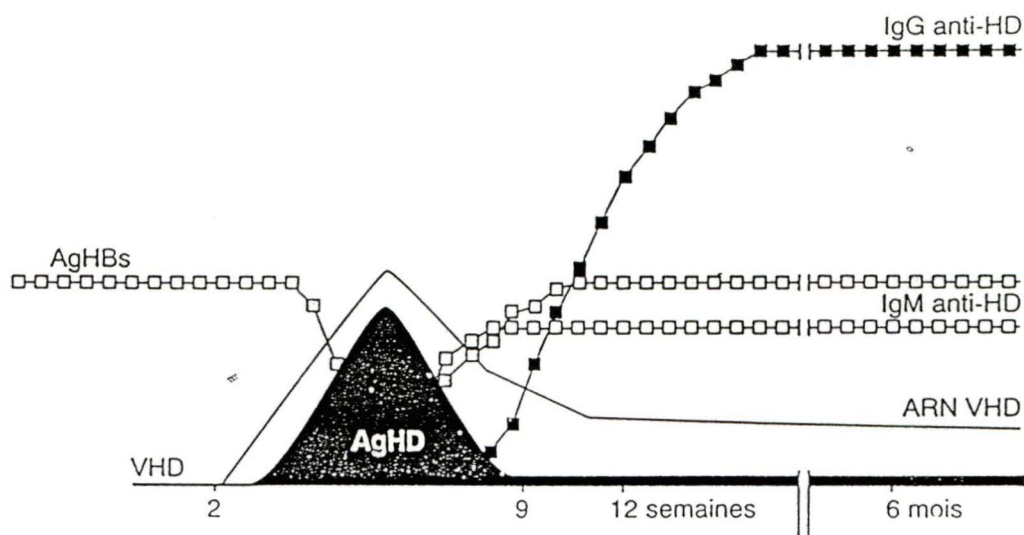


Figure 2. Marqueurs sérologiques d'une surinfection par le VHD d'un porteur chronique du VHB. Lors de la connaissance d'une sérologie VHB, la mise en évidence des marqueurs du VHD permet d'établir le diagnostic de la surinfection *delta* sinon, le diagnostic repose sur la mise en évidence des marqueurs du VHD et de l'AgHBs, en l'absence d'IgM anti-HBc. Au cours de l'infection aiguë, l'AgHD peut être détecté transitoirement dans le sérum par une méthode immuno-enzymatique (ELISA) au cours des premières semaines de l'infection. L'AgHD est rapidement masqué par les IgM et les IgG anti-HD à titre élevé. (Trépo et Valla 1993).

En fait, la recherche des marqueurs d'hépatite *delta* doit être proposée chez tout patient avec une hépatite B aiguë positive pour le marqueur HBs, dans les hépatites B chroniques et chez les populations à risques (pourtour méditerranéen, toxicomanes ou sujets ayant des expositions parentérales multiples et patients séropositifs pour le virus de l'immunodéficience humaine (VIH)).

Concernant un traitement potentiel de l'infection par le VHD, il existe peu d'études biocliniques évaluant les traitements potentiels de l'hépatite chronique *delta*. En Europe du Nord-Ouest et aux États-Unis, les patients atteints d'hépatite D chronique sont surtout des toxicomanes, difficiles à enrôler dans un processus de suivi bioclinique. D'autre part, ces patients sont fréquemment poly-infectés (VIH, VHC). Les meilleurs essais thérapeutiques furent réalisés en Italie du Sud, où un grand nombre de sujets infectés chroniquement par le VHD et appartenant à la population générale ont été suivis à long terme. Les résultats de ces études montrent une efficacité partielle d'un traitement par l'interféron-alpha. Toutefois, lorsqu'un suivi est pratiqué plusieurs mois après l'arrêt du traitement, on observe une réapparition de la virémie chez la majorité des patients. Certaines études suggèrent que l'utilisation de doses plus élevées que celles utilisées pour les infections par le VHB seul pourraient être bénéfiques pour contrer les infections chroniques par le VHD. Or, l'interféron-alpha est très dispendieux. Curieusement, dans les cas de rémission, celle-ci ne corrèle pas toujours avec une disparition de l'ARN viral du VHD dans le sérum des patients. Ceci pourrait suggérer qu'il y a une sélection par le traitement de variants du VHD non cytopathogènes. Enfin, la transplantation hépatique donne des résultats satisfaisants.

La prévention du VHD est identique à celle du VHB. La vaccination contre le VHB, qui neutralise les interactions entre l'enveloppe du VHB (et donc du VHD) et l'hépatocyte, protège de la co-infection VHB/VHD. Pour un sujet porteur chronique du VHB, la protection contre la surinfection *delta* ne repose que sur des mesures de prévention. La mise au point d'une vaccination spécifique contre le VHD n'est pas envisageable. Les modèles animaux ont montré que les anticorps anti-*delta* ne protègent pas de l'infection par le VHD.

1.2.3 Épidémiologie

Mondialement, il y aurait 15 millions d'individus porteurs du virus de l'hépatite *delta* (VHD) et au moins 300 millions porteurs du virus de l'hépatite B (VHB). Du fait que l'enveloppe du VHD soit produite par l'apport de l'AgHBs du VHB, l'épidémiologie et les modes de transmission du VHD sont similaires à celles du VHB. En Europe du Nord et de l'Ouest et aux États-Unis, où l'infection chronique par le VHB est peu fréquente (<1%), l'infection par le VHD est essentiellement confinée aux toxicomanes utilisateurs de drogues par voie parentérale. En revanche, le VHD semble infecter rarement les homosexuels porteurs chroniques du VHB, ce qui pourrait indiquer une transmission sexuelle moins efficace que pour le VHB. Sur le pourtour méditerranéen et au Moyen-Orient, où la fréquence du portage chronique du VHB est intermédiaire (1 à 5%), l'infection par le VHD concerne 5 à 20% de sujets infectés chroniquement par le VHB. Dans ces régions, la transmission intrafamiliale du VHD a été montrée. Actuellement, il semble que l'épidémiologie du VHD dans ces régions soit en train de se

modifier. Par exemple, en Grèce, l'infection par le VHD devient moins fréquente dans la population générale rurale, mais devient élevée dans la population croissante de toxicomanes. En Asie où le portage chronique du VHB est fréquent (10 à 20%), l'infection par le VHD est en revanche très peu rencontrée. En Afrique, où la fréquence du portage chronique du VHB est similaire à l'Asie, ainsi qu'en Amérique du Sud et dans les îles du Pacifique, où elle y est intermédiaire (1 à 5%), l'infection par le VHD est de fréquence extrêmement variable d'une région à une autre. Aussi, des prévalences du VHD de moins de 1% ou de plus de 50% ont été observées à quelques kilomètres de distance, malgré une prévalence du portage chronique du VHB uniforme. Dans les régions où l'endémicité est intermédiaire à élevée, comme en Italie, en Afrique et en Amérique du Sud, la transmission se ferait principalement par contact entre les individus. Par contre, dans les régions où elle est faible, comme en Amérique du Nord, la transmission serait limitée aux groupes d'individus qui ont des expositions parentérales fréquentes comme par exemple les hémophiles et les utilisateurs de drogues injectables. Les variations observées à l'échelle mondiale ou régionale restent inexpliquées, mais elles semblent simplement refléter des différences de pénétration de l'infection dans diverses populations, plutôt que des résistances à l'infection par le VHD (revue par Polish et al., 1993).

1.3 CARACTÉRISTIQUES DU VIRUS DE L'HÉPATITE *DELTA* HUMAINE

Les caractéristiques du VHD le distinguent de tous les virus du règne animal et l'apparentent à des petits ARN pathogènes des plantes, c'est-à-dire les viroïdes. Comme ces derniers, il possède un court génome constitué d'ARN à forte structure secondaire et a recours à un mécanisme en cercle roulant pour effectuer sa réplication. Le motif autocatalytique d'ARN retrouvé dans le génome viral du VHD effectue le clivage en *cis* des brins multimériques produits lors de la réplication afin de libérer les monomères. Le génome du VHD code pour une seule protéine, l'antigène *delta* (AgHD), dont le rôle est essentiel durant la régulation du cycle de réplication. L'AgHD intervient aussi dans l'interaction entre le VHD et le VHB et dans la pathogenèse de la co-infection. Ce virus particulier pourrait être un prototype de nouveaux agents infectieux, tant pour les animaux que pour les hommes (revue par Poisson et al., 1995).

1.3.1 Structure du virion

Le virion complet infectieux du VHD est d'une forme sphérique et à un diamètre de 36 nm. Il comprend l'ARN viral complexé à une vingtaine de multimères d'AgHD et contient, en outre, quelques dizaines de molécules d'AgHD non liées à l'acide nucléique. L'AgHD ne forme pas une capside structurée comme dans les virus conventionnels. Une molécule d'ARN viral et environ 70 molécules d'AgHD-S sont enveloppées dans une

bicouche lipidique contenant l'antigène de surface (AgHBs) du VHB (revue par Poisson et al., 1995). (Figure 3).

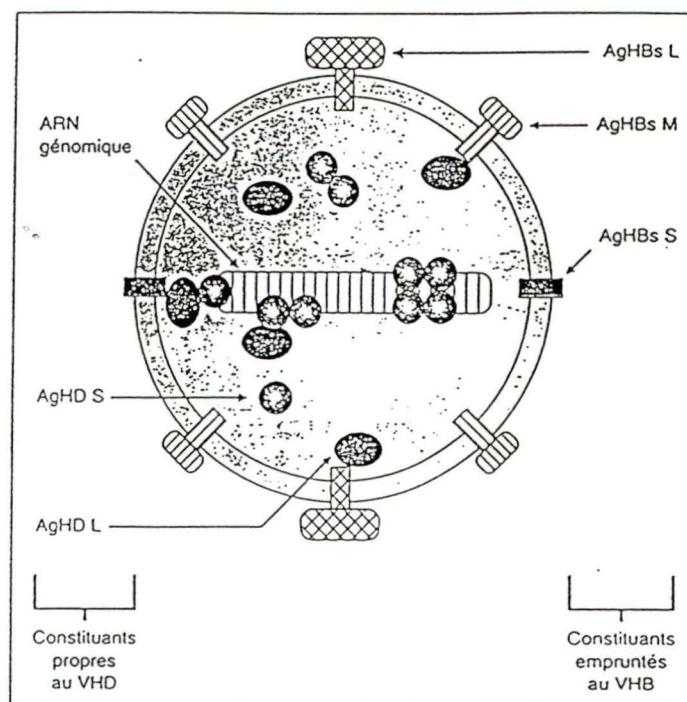


Figure 3: Structure de la particule virale *delta*. Les AgHBs, M et L sont les protéines du VHB formant l'enveloppe du virus de l'hépatite D. Les AgHD-S et L sont les antigènes *delta*. (Poisson et al., 1995).

1.3.2 Le génome viral et les génotypes

Le génome du VHD est un ARN monocaténaire circulaire de 1.7 kb, avec une composition G+C de l'ordre de 60%. De nombreuses complémentarités, à l'intérieur de la séquence du génome viral, permettent au VHD d'adopter une structure secondaire semblable à une tige double-brin (revue par Lai, 1995) (Figure 4). Le génome peut être séparé en deux domaines distincts: le domaine de droite, qui est codant et le domaine de gauche, qui s'apparente fortement aux viroïdes. Le domaine de droite comporte un cadre de lecture ouvert. Celui-ci code pour un seul ARNm d'environ 700 nt, qui peut cependant être traduit à partir de la séquence antigénomique en deux isoformes de l'antigène *delta* (AgHD) (revue par Lazinski et Taylor, 1994). Le domaine terminal gauche porte, quant à lui, tous les éléments nécessaires à l'exécution de la réplication en cercle roulant.

L'infection par le VHD existe à travers le monde. Jusqu'à ce jour, la séquence nucléotidique complète de 22 isolats du VHD de différentes régions a été déterminée. Les analyses d'homologie de séquences révèlent l'existence d'une variabilité entre les isolats du VHD et ceux-ci peuvent être classés en trois génotypes. Cette variabilité entre différents isolats définit le concept d'hétérogénéité de séquences virales. Les isolats appartenant au génotypes II et III présentent respectivement 80% et 65% d'homologie avec ceux appartenant au génotype I. Les isolats du génotype I, présentant plus de 85% d'homologie entre eux, furent retrouvés dans différents pays dont l'Italie, les États-Unis, Taiwan, la France, le Liban et la Chine. Les effets cliniques associés à une infection par

un isolat du VHD de génotype I sont très variables allant de la forme asymptomatique à une forme d'hépatite sévère. Les isolats du génotype II ont été retrouvés jusqu'à maintenant seulement au Japon et à Taiwan. Très peu de symptômes sont associés à une infection causée par un isolat de génotype II. Il s'agit de la forme la plus douce de la maladie. Les isolats du génotype III ont été retrouvés dans les régions au Nord de l'Amérique du Sud dont au Pérou et en Colombie. Les effets cliniques associés à une infection par un isolat du génotype III sont très sévères. La possibilité d'une relation entre l'hétérogénéité de séquence de l'ARN du VHD et les différentes manifestations cliniques observées chez les patients infectés n'est pas très bien comprise.

Il existe un autre niveau de variabilité de séquences de l'ARN viral du VHD. Chez un patients infecté par le VHD, l'ARN se retrouve sous forme d'une population hétérogène d'ARN viral (quasi-espèce) ce qui définit la microhétérogénéité (revue par Poisson et al., 1999). Enfin, l'ARN du VHD serait en continuelle évolution à travers l'infection. Il est suggéré que le taux d'évolution approximatif est de 3×10^{-2} à 3×10^{-3} substitutions de nucléotides par année. Le taux d'évolution varie selon le domaine du génome d'ARN viral et le stade clinique de l'infection (revue par Lai, 1995).

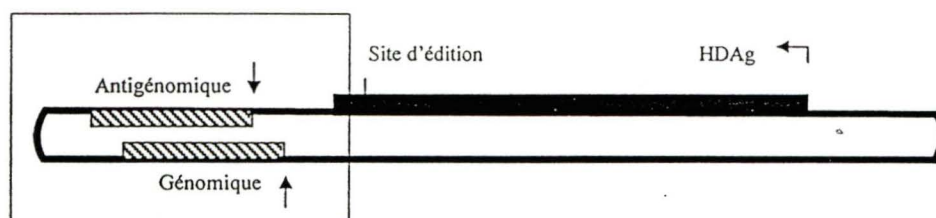


Figure 4: Représentation schématique de la structure secondaire en tige du génome du virus de l'hépatite *delta*. Le domaine terminal de gauche qui est encadré possède les caractéristiques retrouvées chez certains viroïdes. Les rectangles lignés correspondent aux séquences des motifs autocatalytiques *delta* génomique et antigénomique et les flèches indiquent le site de clivage. Le rectangle noir représente le cadre de lecture encodant les deux formes d'antigènes *delta* résultant ou non d'un événement d'édition.

1.3.3 Pathogenèse

Les mécanismes impliqués dans la pathogenèse du VHD sont encore mal connus. Des études histologiques ont suggéré un effet cytopathique direct, tandis que d'autres ont suggéré un mécanisme indirect, via une réponse immunitaire cellulaire T, comme pour le VHB (Trépo et Valla, 1993). Un effet cytopathique direct du VHD pourrait s'exercer par plusieurs mécanismes tels que: 1) une cytotoxicité directe de l'AgHD-S exprimée dans la cellule infectée; 2) une compétition entre l'ADN cellulaire et l'ARN du VHD pour l'ARN polymérase II; 3) l'existence d'une complémentarité entre l'ARN antigénomique du VHD et un ARN cellulaire, impliquant la formation d'un hybride et l'altération des fonctions cellulaires. Il est possible que ces trois types de mécanismes directs, ainsi qu'une réponse immunitaire, se superposent pour induire un effet cytopathogène. De plus, la proportion du génome codant pour la forme S versus la forme L de l'AgHD pourrait aussi intervenir dans le degré de cette pathogénicité.

Il faut noter que le séquençage de plusieurs isolats du VHD dans différentes régions géographiques a montré des variations de séquence importantes. La variabilité observée dans la séquence virale du VHD pourrait être impliquée dans les différences de pathogénicité des isolats, expliquant la tolérance de certaines populations pour le VHD. De plus, il y aurait possiblement des variations se produisant dans le temps, au cours de l'infection chez un sujet. Or, les relations entre géotypes, degré de variabilité des

séquences, distribution géographique et sévérité de la maladie sont des aspects encore incompris.

1.3.4 Mécanisme de réplication virale

Le cycle de vie du VHD débute par l'adhérence et la pénétration du virion dans l'hépatocyte. Le complexe ribonucléoprotéique formé par l'ARN viral complexé à des AgHD est transporté du cytoplasme vers le noyau, par l'intermédiaire des AgHD qui comportent un signal de localisation nucléaire. Ce transport est très important car l'enzyme responsable de la réplication virale est localisée au noyau. Le cycle de réplication de l'ARN du VHD est mis en route par l'ARN polymérase II de l'hôte (qui normalement utilise un ADN double-brin comme matrice) (revue par Poisson et al., 1995). Il est suggéré que la structure double-brin de l'ARN du VHD permettrait de tromper la machinerie cellulaire. Dans l'hépatocyte infecté, les ARN génomiques et antigénomiques sont détectés sous forme linéaire, circulaire ainsi que sous forme de multimères. La présence de ces multimères s'explique par une réplication dite "en cercle roulant" (Figure 5). Ce mécanisme de réplication est complètement indépendant d'ADN (revue par Lazinski et Taylor, 1995). Au cours de la première phase d'initiation, l'ARN polymérase II reconnaît un promoteur sur le brin génomique permettant la synthèse d'un brin multimère antigénomique, c'est-à-dire de polarité opposée. Ce nouveau brin est coupé en copies monomériques qui sont liguées ensuite pour produire des copies circulaires antigénomiques. Par la suite, ces molécules d'ARN circulaires servent de

matrice pour la production de la progéniture circulaire génomique, qui s'effectue selon les trois mêmes étapes de synthèse, clivage et ligation. Ce mécanisme est extrêmement efficace puisqu'une cellule infectée peut contenir jusqu'à 300000 copies d'ARN génomique du VHD. Le clivage des intermédiaires multimériques de réplication est accompli par deux domaines autocatalytiques distincts, situés sur l'ARN viral du VHD, qui ont respectivement le rôle de couper les brins génomique et antigénomique. Ces domaines, nommés motifs autocatalytiques *delta*, agissent en *cis*, c'est-à-dire qu'ils permettent l'autoclivage de leur propre brin d'ARN à un site spécifique (revue par Mercure et al., 1997).

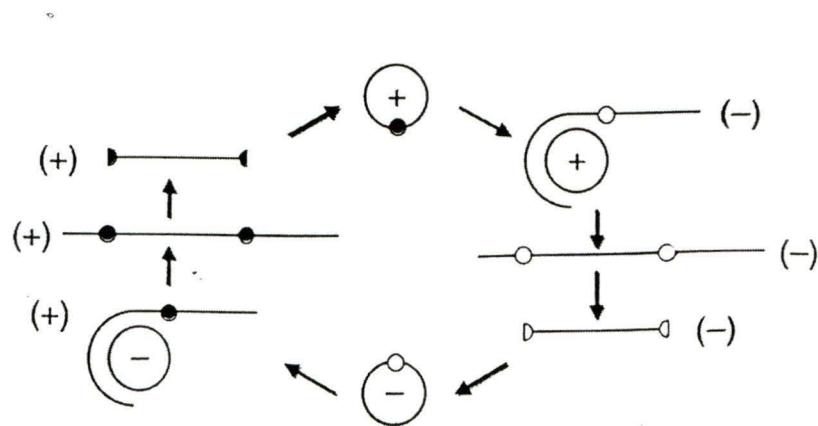


Figure 5: Représentation du cycle de réplication en cercle roulant du VHD. Les petits cercles noirs et blancs représentent les sites d'autocoupeure des brins génomique et antigénomique. Les polarités génomique et antigénomique sont représentées respectivement par les symboles (+) et (-). (Mercure et al., 1997).

1.3.4.1 Les domaines autocatalytiques (ribozymes *delta*)

Tout comme chez certains viroïdes, on retrouve au sein du génome du VHD de courts domaines d'ARN, de moins de 100 nucléotides, qui sont indispensables à la réplication de son génome d'ARN. Ces domaines se caractérisent par leur capacité d'effectuer le clivage des intermédiaires de réplication du génome. Chez le VHD, ce domaine est connu sous le nom de motif d'autocoupure *delta* (Been, 1994). L'analyse de fragments d'ARN d'origine génomique et antigénomique révéla la présence d'un site d'autocoupure spécifique sur chacun de ces brins. Bien que la séquence primaire de ces domaines catalytiques soit relativement différente, les structures secondaires prédites s'avèrent presque identiques d'où l'appellation commune, «ribozyme *delta*», pour ces deux entités. Par suite de l'appariement partiel créant la structure en tige du génome viral, les domaines autocatalytiques génomique et antigénomique se retrouvent directement opposés l'un à l'autre (Figure 4). Lors de la réplication du VHD, le mécanisme de clivage des brins génomiques et antigénomiques est dit autocatalytique puisque la réaction a lieu *en cis* et une seule fois. Toutefois, il a été possible de diviser les brins d'ARN génomique et antigénomique en deux domaines et de développer une réaction *en trans* dans laquelle une molécule, identifiée comme un ribozyme, possède les propriétés catalytiques nécessaires au clivage successif de plusieurs molécules de substrats (Figure 6) (revue par Mercure et al., 1997).

Les ribozymes peuvent être considérés comme des agents thérapeutiques prometteurs pour moduler des ARN donnés dans le contexte cellulaire. Par exemple, ils

pourraient cibler puis cliver des ARN viraux de façon à prévenir la réplication ultérieurs de ces pathogènes, ou encore moduler le produit de l'expression de gènes tels les oncogènes impliqués dans différents cancers humains.

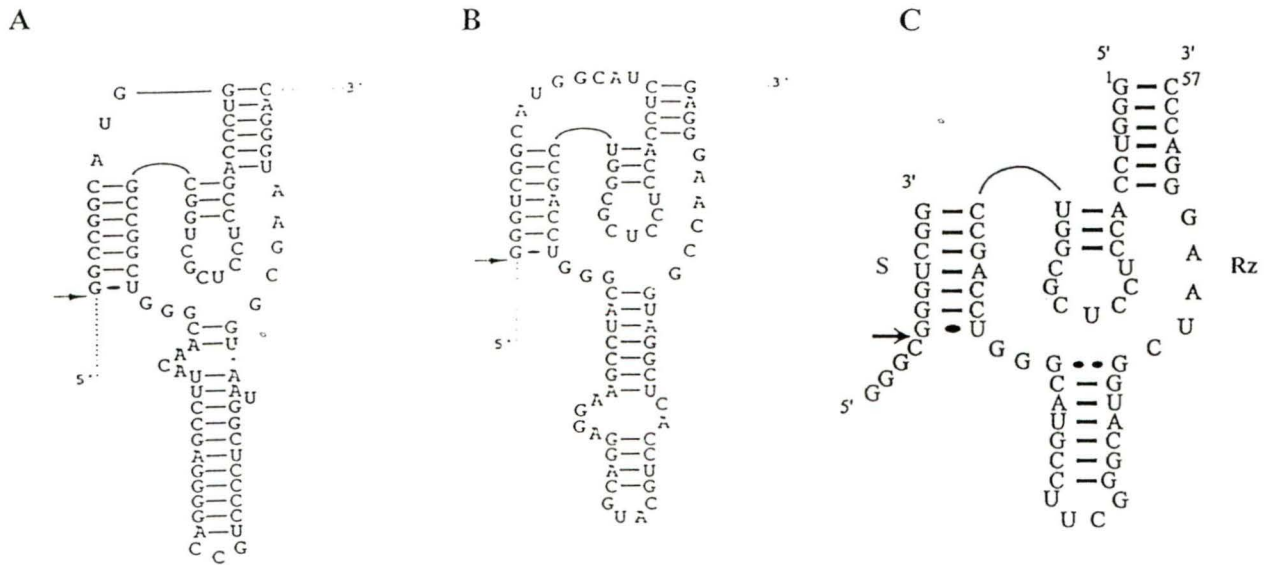


Figure 6: Représentation schématique des motifs autocatalytiques d'ARN du VHD. Séquences nucléotidiques et structures secondaires des motifs autocatalytiques génomique (A) et antigénomique (B) dans la réaction en *cis*. Construction du ribozyme agissant en *trans* (C). Celle-ci correspond à celle que nous utilisons et est dérivée de la séquence antigénomique. Les flèches indiquent les sites de coupure.

1.3.4.1 Les antigènes *delta*

L'antigène *delta* est soit d'une forme petite de 195 acides aminés (a.a.) (AgHD-S) ou d'une forme longue de 214 a.a. (AgHD-L). La différence entre ces deux isoformes se retrouve dans la portion C-terminale. La forme allongée résulte d'un événement d'édition au niveau de l'ARNm. De récents résultats démontrent qu'une désadénylation transforme le codon de terminaison du cadre de lecture de la forme courte en un codon tryptophane, ce qui permet l'élongation du cadre de lecture ouvert de 19 acides aminés (a.a.) (Polson et al., 1996). Le prochain codon de terminaison est alors reconnu comme signal d'arrêt de la traduction. Quoique très semblables au niveau de leurs contenus en acides aminés, les deux formes d'antigènes ont des rôles biologiques très différents. La forme de 195 acides aminés est, entre autre, nécessaire pour permettre le début de la réplication tandis que la forme de 214 acides aminés favorise les étapes suivantes d'assemblage de la capside aux dépends de celles de réplication (revue par Poisson et al., 1995).

1.3.5 Taxonomie

Les caractéristiques du virus de l'hépatite *delta* humaine (VHD) le distinguent de tous les virus du règne animal. Certains le classifie comme une particule sous-virale, un virus défectif ou encore un virus satellite puisqu'il nécessite la présence d'un autre virus (VHB) pour se propager (revue par Hadziyannis, 1997; Carreno et al., 1994; Lai, 1995; Lazinski et Taylor, 1994). Le VHD se distingue également des autres virus puisqu'il

possède un court génome d'ARN monocaténaire circulaire comprenant un fort degré d'appariement interne. Ceci le conduit à adopter une structure secondaire en forme de longue tige double-brin. De plus, la réplication est indépendante de l'ADN et s'effectue selon un mécanisme en cercle roulant. Bien qu'aucune homologie de séquence puisse être déduite, le VHD possède plusieurs similarités avec les viroïdes et les ARN satellites de virus qui infectent les plantes (Diener, 1993). Tout comme dans certains viroïdes, on retrouve au sein du génome du VHD de courts domaines d'ARN autocatalytiques, de moins de 100 nucléotides, qui sont indispensables à la réplication de son ARN. Toutefois, contrairement aux viroïdes, le VHD possède un cadre de lecture ouvert codant pour une protéine, l'antigène *delta*. Enfin, le VHD s'apparente aux viroïdes par son autonomie de réplication. En effet, à l'opposé des ARN et des virus satellites qui utilisent les enzymes de réplication de leurs virus auxiliaires, le VHD se réplique sans aucun facteur du VHB, dont le rôle se limite à fournir une enveloppe au génome du VHD (revue par Poisson et al., 1997). Le virus de l'hépatite *delta* appartient donc à une nouvelle classe d'agents infectieux, jusqu'alors insoupçonnée chez l'animal et dont le groupe taxonomique attribué est *Deltaviridae*. Le tableau 1 met en évidence les principales similitudes et différences entre les viroïdes et le VHD.

Tableau 1: Similitudes et différences entre le VHD et les viroïdes.

Caractéristiques	Viroïdes	VHD
Acide nucléique du génome	ARN sb	ARN sb
Taille du génome	246-475 nt	1650-1700 nt
Forme du génome	Circulaire	Circulaire
Mécanisme de réplication	Cercle roulant	Cercle roulant
Structure en tige double-brin	Oui	Oui
Hôte	Plantes	Humain
Code pour des protéines	Non	Oui

1.4 APERÇU ET OBJECTIFS DE LA RECHERCHE

Les infections virales par le VHD et le VHB constituent un problème de santé publique important. À ce jour, il n'existe aucun traitement efficace pour venir en aide aux sujets atteints d'une hépatite infectieuse de type D d'où l'importance de poursuivre les recherches dans ce domaine. En raison de la faible efficacité et des coûts exubérants reliés à son utilisation, l'interféron-alpha ne semble pas être la solution pour vaincre l'infection par le VHD. De plus, la diversité génétique due à la variabilité des séquences virales du VHD pourrait possiblement expliquer en partie pourquoi le VHD réussit à déjouer la surveillance immunitaire de l'hôte, a un haut taux de chronicité, échappe souvent au traitement et rend le développement d'un vaccin presque impensable. De plus, les caractéristiques biochimiques du VHD en font un virus unique très intéressant. Plusieurs questions demeurent sans réponse sur différents aspects du VHD. Le développement d'outils de recherche visant à faciliter l'étude et la compréhension des propriétés biologiques du VHD est au coeur de notre projet de recherche.

L'objectif général du projet de recherche exposé consiste à approfondir nos connaissances de l'aspect moléculaire du VHD et de développer des outils pouvant être utiles pour les travaux visant, entre autre, l'inhibition de la propagation de ce pathogène humain. L'approche utilisée pour parvenir à ces fins comporte trois étapes. Dans un premier temps, de façon à contribuer à l'étude des mécanismes d'évolution et d'approfondir nos connaissances de l'aspect moléculaire du VHD, nous allons procéder à

l'alignement précis des séquences nucléotidiques complètes connues du VHD ainsi qu'à l'analyse de cet alignement. L'alignement sera mis à la disposition de la communauté scientifique dans une banque de données dont je participerai à la réalisation. Celle-ci pourra servir d'outil par exemple, pour l'étude des concepts d'hétérogénéité, de microhétérogénéité et de quasi-espèce, relatifs au VHD ou autres ARNs apparentés.

En deuxième lieu, l'établissement d'une relation entre les génotypes, le degré de variabilité de séquences, la distribution géographique et le niveau de la sévérité de la maladie nécessite l'analyse d'un grand nombre de séquences virales du VHD. C'est pourquoi, nous allons déterminer la séquence nucléotidique complète du VHD isolé d'un sujet canadien infecté par ce virus. L'ADN complémentaire (ADNc) de l'ARN viral extrait sera produit par transcription inverse (RT) et amplifié par une réaction de polymérisation en chaîne (PCR). De façon à optimiser les réactions de synthèse et d'amplification de l'ADNc, il sera primordial de déterminer des amorces optimales pouvant amplifier n'importe quelle séquence virale. En effet, il y a beaucoup de variations de séquence au sein du génome du VHD. Pour parvenir à cette fin, nous allons utiliser l'alignement des séquences du VHD de façon à déterminer des amorces situées dans des régions conservées du génome viral, chez les trois génotypes. Après clonage, le séquençage de quatre clones indépendants permettra de vérifier le concept de microhétérogénéité des séquences virales au sein d'un sujet infecté.

Dans un troisième temps, l'étude des caractéristiques biochimiques que confèrent les motifs autocatalytiques du VHD de même que les travaux visant l'inhibition de la

propagation de ce pathogène humain nécessitent un outil de travail adéquat. C'est pourquoi nous allons développer une nouvelle lignée cellulaire hépatique (HepG2) exprimant constitutivement l'ARNm correspondant à la forme petite de l'antigène *delta* (AgHD-S). Cette lignée sera établie par la transfection stable d'un vecteur d'expression, que nous possédons au laboratoire, dans lequel un insert correspondant à l'ADNc de l'AgHD-S a été cloné. La caractérisation des colonies cellulaires issues de la transfection stable sera réalisée par la recherche de celles qui expriment l'ARNm AgHD-S de façon constitutive. L'identification des colonies positives sera faite par RT-PCR et par buvardage de type Southern. En bref, ce projet de recherche repose sur le développement d'outils nécessaires pour la compréhension de différents aspects du virus de l'hépatite *delta* humaine.

2.0 MATÉRIEL ET MÉTHODES

2.1 ALIGNEMENT ET ANALYSE DES SÉQUENCES COMPLÈTES DU VHD

Le laboratoire a conçu une banque de séquences qui a été mise à la disposition de la communauté scientifique. L'objectif relié à la réalisation de cette banque est de faciliter les recherches sur les viroïdes et autres ARN apparentés dont le virus de l'hépatite *delta* humaine en fournissant de l'information relative aux aspects moléculaires de ces espèces d'ARN. Cette banque d'information est disponible par Internet à l'adresse suivante: <http://www.callisto.si.usherb.ca/~jpperra> (Lafontaine et al. 1999).

Toutes les séquences nucléotidiques complètes du VHD connues et répertoriées dans la banque de données du laboratoire furent utilisées pour cette étude. Les 22 séquences de polarité génomique utilisées ont d'abord été alignées à l'aide du programme informatique "Clustal W" (version 1.6). Ensuite, de façon à optimiser l'alignement des séquences, nous avons procédé à la vérification systématique de l'alignement complet manuellement. Enfin, diverses informations relatives aux caractéristiques du génome viral du VHD ont été recherchées par l'analyse des séquences alignées.

2.2 SÉRUM D'UN PORTEUR CANADIEN

Cette étude a été approuvée par le groupe d'éthique du Centre de Recherche Clinique de l'Université de Sherbrooke. Pour ce projet, le sérum d'un seul sujet infecté par le VHD fut utilisé. L'échantillon de sérum a été prélevé d'un sujet mâle, âgé de 50 ans et résidant au Canada dans la province de Québec, avec la collaboration du Dr Daniel B. Ménard, gastroentérologue. L'histoire clinique de ce patient établit qu'il a souffert d'une hépatite chronique B sévère qui a rapidement évoluée vers une cirrhose du foie. Les cliniciens avaient procédé à la caractérisation sérologique de ce patient. En fait, le sérum fut analysé pour les antigènes de surfaces du virus de l'hépatite B (HBs) et les IgM anti-HBc. Ensuite, le diagnostic de l'infection par le VHD fut établie par la détection des antigènes totaux du VHD et les IgM anti-HD. Enfin, les anticorps relatifs à une infection par le VHC furent détectés.

2.3 ARN VIRAL DU VIRUS DE L'HÉPATITE *DELTA* HUMAINE

2.3.1 Extraction de l'ARN viral

L'extraction de l'ARN viral du VHD a été réalisée avec la collaboration du Dr Eric Frost du département de Microbiologie et d'infectiologie du Centre Universitaire de Santé de l'Estrie. L'ARN a été extrait à partir du sérum du porteur canadien par une

digestion à la protéinase K suivit d'une extraction phénol/chloroforme (v/v). L'étape d'extraction a été suivie d'une précipitation à l'isopropanol (0.7v) de la phase aqueuse et de lavages à l'éthanol 70%. L'ARN du VHD a été ensuite resuspendu dans un volume de 50 µl d'eau préalablement traité avec du DEPC (diéthyl pyrocarbonate) 0,1% incluant 30 unités d'inhibiteur de ribonucléase (RNAguard™, Amersham Pharmacia Biotech Inc.).

2.3.2 Détermination et synthèse des amorces RT-PCR

La détermination des amorces nécessaires pour les réactions de transcription inverse (RT) et de polymérisation en chaîne (PCR) s'est fait à l'aide de l'alignement des séquences nucléotidiques complètes du VHD. D'abord, toutes les amorces ont été sélectionnées dans des régions hautement conservées du génome viral. La séquence nucléotidique de chacune des amorces correspond à celle du variant vHDV.5 (Lafontaine et al., 1997) et est très bien conservée dans les 22 séquences complètes connues du VHD. Donc, la séquence des amorces est conservée pour chacun des trois génotypes. La numérotation des amorces correspond à celle de la séquence originale déterminée par Makino et al., 1987. Les oligonucléotides servant d'amorces ont été synthétisés par la compagnie Gibco BRL (Burlington, Ontario).

La vérification systématique de chacune des amorces pour leur capacité à s'hybrider et à amplifier seulement la séquence du VHD a été faite de deux façons. D'abord, à l'aide du programme informatique "Amplify" et ensuite à l'aide d'un serveur

disponible au site du «National Center for Biotechnology Information» (NCBI) à l'adresse Internet suivante: <http://www.ncbi.nlm.nih.gov/>

2.3.3 Synthèse de l'ADNc

La production de l'ADNc s'est faite à partir de l'ARN viral du VHD extrait du sérum du porteur canadien par une technique de transcription inverse. L'oligonucléotide utilisé (5'dGGAAAAGAAGAGTAGCCG3'), nommée AS1100, correspond à la position 1163 à 1180 du variant vHDV.5 (Lafontaine et al., 1997) et a été déterminé à l'aide de l'alignement des séquences nucléotidiques complètes du VHD. La réaction de transcription inverse a été réalisée à partir de 1/10 (5 µl) du volume de l'ARN viral extrait. Une quantité de 100 pmol d'amorce AS1100 fut utilisée dans un volume de 18 µl. De façon à optimiser l'efficacité de la réaction de RT, une incubation à une température de 70°C pour une durée de 5 min a été faite pour permettre la dénaturation des structures secondaires pouvant être présentes dans l'ARN. Ensuite, le volume du mélange réactionnel a été ajusté à 30 µl par l'ajout de 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM dithiothréitol (DTT) 0,67 mM de chacun des dNTP (dATP, dCTP, dGTP, dTTP), 30 unités d'inhibiteur de ribonucléase (RNAGuard™, Amersham Pharmacia Biotech Inc.) et 200 unités de la transcriptase inverse (SuperScript™ II, Gibco BRL). Suite à une incubation à 42°C d'une durée de 50 minutes, la réaction fut arrêtée par l'inactivation de l'enzyme à l'aide d'une incubation de 15 minutes à 70°C suivit de 15 minutes sur glace.

2.3.4 Amplification de différents fragments d'ADNc

L'amplification de l'ADNc a été faite par la technique de polymérisation en chaîne (PCR) en utilisant quatre paires d'amorces (voir Résultats, section 3.3.1). Quatre fragments d'ADNc qui se chevauchent furent synthétisés par PCR. Les quatre réactions de PCR ont été réalisées dans un volume de 100 µl à partir de 1/5 (6 µl) de la réaction de RT. Le mélange réactionnel contenait 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM MgCl₂, 0.2 mM de chacun des dNTP, 100 pmol de chacune des amorces (sens et antisens) et 1 unité de l'ADN polymérase *Taq* purifiée dans notre laboratoire selon le protocole de Engelke et al. (1990). Trente cycles d'amplification (1 min à 94°C, 1.5 min à une température entre 52°C et 65°C selon la paire d'amorces utilisée, 1 min à 72°C) ont été réalisés. Finalement, une extension terminale de 10 min à 72°C a été faite.

2.3.5 Purification des différents fragments d'ADNc

Suite à l'amplification des quatre différents fragments d'ADNc, le mélange réactionnel a été précipité à l'isopropanol suivit d'étapes de lavages à l'éthanol 70%. Un aliquot de 20 µl des produits de PCR précipités et resuspendus a été séparé par électrophorèse sur gel d'agarose 1.5%. Les bandes de gel contenant chacun des quatre produits de PCR ont été ensuite identifiées par coloration au bromure d'éthidium, découpées et l'ADN a été purifié à l'aide de la trousse "QIAquick gel extraction kit" (QIAGEN).

2.3.6 Clonage des différents fragments d'ADNc

Chacun des quatre fragments d'ADNc a été cloné séparément en utilisant la propriété de l'ADN polymérase *Taq* d'ajouter une désoxyriboadénosine (dATP) à l'extrémité 3' du produit de PCR. Les produits d'amplification PCR ont été ligués dans le vecteur linéaire pCR™ 2.1 qui possède un résidu désoxythymidine (dTTP) à chacune des extrémités 5', suivant les indications du manufacturier (TA cloning kit, Invitrogen). Les réactions de ligation ont été réalisées à une température de 14°C pour une durée de 16 heures dans un tampon et une enzyme ADN-ligase fournies par le manufacturier. Les plasmides obtenus ont finalement été amplifiés par des bactéries compétentes (Top10F') également fournies par le manufacturier. La transformation s'est déroulée de la façon suivante: 2 µl de la réaction de ligation et 2 µl de β-mercaptoéthanol 0.5 M ont été ajoutés à 50 µl de cellules compétentes. Une incubation sur la glace d'une durée de 30 min et ensuite à 42°C pour 30 s suivit d'une réincubation sur glace pour 2 min a été faite. 200 µl de milieu de culture SOC (bacto-tryptone 20 g/l, bacto yeast-extract 5 g/l, NaCl 0.5 g/l, glucose 20 mM) pré-chauffé à 37°C ont été ajoutés. Le tout a été incubé à 37°C durant 1 heure. Le mélange a été ensuite étalé sur des pétris composés de milieu LB (bacto-tryptone 10 g/l, bacto yeast-extract 5 g/l, NaCl 10 g/l), 1.5% agar, 10 µg/ml ampiciline) qui sont recouverts d'une solution contenant 40 µl de Xgal (isopropyl-β-D-thiogalactopyranoside) à 20 mg/ml et 4 µl d'IPTG (5-bromo-4-chloro-3-indolyl-β-D-

galactopyranoside) à 200 mg/ml. Les plasmides ayant transformé les bactéries ont été isolés par mini-préparation selon la méthode de lyse alcaline (Sambrook et al., 1989).

2.3.7 Séquençage des différents fragments d'ADNc

Quatre colonies positives isolées par mini-préparation (clones) correspondant à chacun des produits de PCR furent conservées. Les plasmides construits ont été séquencés par la technique décrite par Sanger (Sanger et al., 1977). Le "T7 Sequencing™ kit" (Amersham Pharmacia Biotech Inc.) fut utilisé avec 20 $\mu\text{Ci/ml}$ [α - ^{35}S -dATP (Amersham Canada Ltd., Oakville, Canada). Les réactions de séquence ont été réalisées avec l'amorce universelle M13 ou avec d'autres amorces identifiées à l'aide de l'alignement de séquences complètes du VHD. La séquence nucléotidique des amorces de séquençage ainsi que leur position au sein du génome viral du VHD sont montrées dans le tableau 2 (voir Résultats, section 3.3.1). Ces amorces furent nécessaires pour le séquençage des régions éloignées des extrémités de certains ADNc relativement longs. Les échantillons ont été par la suite séparés par électrophorèse sur un gel dénaturant composé d'urée-polyacrylamide 5% (voir Annexes, section 8.1). Le gel fut ensuite séché sous vide durant 30 min. à 80°C et les bandes révélées par autoradiographie. Au fur et à mesure que la séquence de chacun des fragments d'ADNc était déterminée, celle-ci était ajoutée à l'alignement des séquences du VHD.

2.4 DÉVELOPPEMENT D'UNE LIGNÉE CELLULAIRE HepG2

2.4.1 Mise en culture des cellules

Pour cette étude, nous avons utilisé les cellules humaines HepG2, dérivées d'un hépatoblastome (ATCC HB-8065). Elles ont été généreusement fournies par le Dr Robert Tanguay du laboratoire de Génétique Cellulaire et Développementale de l'Université Laval. Ces cellules provenaient de l'Américain Type Culture Collection (Rockville, MD). Les cellules HepG2 ont été maintenues dans un incubateur à une température de 37°C en présence de 5% CO₂ et 95% air. Le milieu utilisé est le MEM (*Eagle's Minimal Essential Medium*) (Gibco BRL) supplémenté avec 10% de sérum fœtal bovin (SFB, Gibco BRL) décomplémenté préalablement 30 min à 56°C. De plus, une solution contenant un mélange d'antibiotiques et d'antimycotiques (Gibco BRL), 2 mM L-glutamine (Gibco BRL) et 1 mM sodium pyruvate (Gibco BRL) a été ajoutée au milieu de culture.

2.4.2 Transfection stable des vecteurs d'expression pcDNA₃ et pcDNA₃-Ag-S et isolation des colonies cellulaires.

Le vecteur pcDNA₃-Ag-S qui a été utilisé pour cette étude était déjà disponible dans le laboratoire. Cette construction correspond au vecteur d'expression eucaryote pcDNA₃ (Invitrogen) qui contient l'ADNc correspondant à l'ARN messager (ARNm) de la forme petite de l'antigène *delta* (HDAg-S). La séquence de l'insert correspond aux

positions 900 à 1679 du variant vHDV.5 (Lafontaine et al., 1997). Ce vecteur permet l'expression de l'HDAg-S, dans des cellules en culture, sous le contrôle du promoteur CMV.

La transfection stable du vecteur pcDNA₃ sans insert (contrôle négatif) et de la construction pcDNA₃-Ag-S a été réalisée à l'aide d'une préparation de liposomes fournie dans la trousse "Dosper Liposomal Transfection Reagent" (Boehringer Mannheim) selon les indications du manufacturier. Suite à la transfection, les cellules ont été sélectionnées à l'aide de la Geneticin™ (G-418, Gibco BRL). Les colonies de cellules ayant intégrées le vecteur, c'est-à-dire démontrant une résistance au G-418, ont été isolées à l'aide d'anneaux de clonage et remises en culture.

2.5 CARACTÉRISATION DES COLONIES EXPRIMANT L'ARN AgHD-S

2.5.1 Extraction d'ARN cytoplasmique

L'ARN cytoplasmique a été extrait des colonies de cellules suivant la technique d'extraction décrite par Sambrook et al., 1989. Brièvement, les cellules ont été lavées au PBS (NaCl 0.14 M, KCl 2.7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1.8 mM), recueillies par centrifugation et resuspendues dans 200 µl de tampon d'extraction d'ARN (NaCl 0.1 M, MgCl₂ 0.01 M Tris-HCl, pH 7.9, Nonidet P-40 (NP-40) 0.65%). Suite à une centrifugation, 200 µl de tampon de digestion protéique (Tris/HCl, pH 7.9, Urée 7 M,

NaCl 0.35 M, éthylène diamine tétraacétate (EDTA) 0.01 M, sodium dodécyl sulfate (SDS), 1%) ont été ajoutés au surnageant. Une extraction phénol/chloroforme (v/v) a été suivit d'une précipitation à l'éthanol 100% de la phase aqueuse et de lavages à l'éthanol 70%. L'ARN a été ensuite resuspendu dans un volume de 20 µl d'eau contenant 30 unités d'inhibiteur de ribonucléase (RNAguard™, Amersham Pharmacia Biotech Inc.). La préparation d'ARN a ensuite été traitée avec 2 unités de désoxyribonucléase (RQ1 DNase, Proméga) à 37°C. Une deuxième étape d'extraction phénol/chloroforme (v/v), précipitation à l'isopropanol (0.7v) et lavages à l'éthanol 70% a suivi la digestion à la DNase. L'ARN a finalement été resuspendu dans 20 µl d'eau préalablement traité avec du DEPC 0,1% et de 30 unités d'inhibiteur de ribonucléase.

2.5.2 Synthèse des ADNc respectifs des ARN AgHD-S et GAPDH

La production de l'ADNc s'est faite à partir de l'ARN cytoplasmique extrait des colonies cellulaires par transcription inverse. L'oligonucléotide utilisé pour la synthèse de l'ADNc de l'ARNm AgHD-S (5'dCCAAGCTTCGAAGAGGAAAGAAG3'), nommé HDV900.14 a été déterminé à l'aide de l'alignement et la séquence correspond à celle du variant vHDV.5 (Lafontaine et al., 1997). L'oligonucléotide utilisé pour la synthèse de l'ADNc de l'ARNm GAPDH (5'dGTCATCCATGACAACCTTTGG3'), nommé GAPdown, correspond aux positions 3974-3993 de l'exon 6 du gène de l'enzyme humaine glycéraldéhyde-3-phosphate-déshydrogénase (GAPDH) (Lafontaine et al., 1998). La réaction de transcription inverse a été réalisée à partir de 5 µg d'ARN

cytoplasmique extrait des cellules. Une quantité de 100 pmol d'amorce HDV900.14 ou GAPdown a été utilisée pour la réaction d'hybridation à l'ARN viral dans un volume de 12 µl. Cette réaction d'hybridation a été faite à une température de 70°C pour une durée de 5 min. Ensuite, le volume du mélange réactionnel a été ajusté à 20 µl par l'ajout de 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 0.67 mM de chacun des dNTP, 30 unités d'inhibiteur de ribonucléase et 200 unités de la transcriptase inverse (SuperScript™ II, Gibco BRL). Suite à une incubation à 42°C d'une durée de 50 min, l'enzyme a été inactivée par une incubation à 70°C pendant 15 min suivit de 15 min sur glace.

2.5.3 Amplification des ADNc AgHD-S et GAPDH

L'amplification des ADNc correspondant à l'AgHD-S et au GAPDH a été faite par la technique de polymérisation en chaîne (PCR). Les amorces sens HDV1679.66 (5'dCCGGATCCCTCGGGCTCGGGCG3') et antisens HDV900.14 (5'dCCAAGCTTCGAAGAGGAAAGAAG3') ont été utilisées pour la réaction de PCR à partir de l'ADNc AgHD-S. Ces amorces s'hybrident à chacune des extrémités de l'ADNc HDAg-S permettant l'amplification d'un fragment de 779 pb. Les amorces sens GAPup (5'dGTGAAGGTCGGTGTCAACGG3') et antisens GAPdown (5'dGTCATCCATGACAACCTTTGG3') ont été utilisées pour la réaction de PCR à partir de l'ADNc GAPDH. L'amorce GAPup correspond aux positions 1462-1481 de l'exon 1 du gène de l'enzyme humaine glyceraldéhyde-3-phosphate-déshydrogénase (GAPDH)

(Lafontaine et al., 1998). L'utilisation de ces amorces permet d'obtenir un produit d'amplification de 548 pb. Les réactions de PCR ont été réalisées dans un volume de 100 μ l à partir de la réaction de RT (20 μ l). Le mélange réactionnel contenait 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 mM de chacun des dNTP, 100 pmol de chacune des amorces (sens et antisens) et 1 unité de l'ADN polymérase *Taq* purifiée dans notre laboratoire selon le protocole de Engelke et al. (1990). Un premier cycle d'amplification de faible stringence (5 min à 94°C; 30 s à 53°C, 1 min à 72°C) a été réalisé. Ensuite, 35 cycles d'amplification (1 min à 94°C, 30 s à 62°C, 1 min à 72°C) ont été réalisés. Finalement, une extension terminale de 10 min à 72°C a été faite. Un aliquot de 20 μ l des produits de PCR (AgHD-S et GAPDH) a été séparé par électrophorèse sur deux gels d'agarose 1% différents pour les deux types de produits de PCR. Les bandes de gel correspondant aux produits attendus ont été ensuite identifiées par coloration au bromure d'éthidium.

2.5.4 Buvarpage de type Southern

Une fois l'électrophorèse terminée, une membrane de transfert en nylon (Hybond N+, Amersham) a été placée sur chaque gel et l'ADN fut transféré par capillarité avec un tampon NaOH 0.4 M. Suite au transfert, les membranes ont été cuites sous vide dans un four à 80°C durant 2 heures.

2.5.5 Synthèse *in vitro* et purification de la sonde d'ARN HDAG-S

Le plasmide utilisé pour la synthèse *in vitro* de la sonde, pKSAgS, était déjà disponible dans le laboratoire. Cette construction correspond au plasmide pBluescript KS+ (Stratagen) dont l'insert est l'ADNc correspondant à l'ARNm de la forme petite de l'antigène *delta* (HDAG-S). La séquence de l'insert correspond aux positions 900 à 1679 du variant vHDV.5 (Lafontaine et al., 1997). La transcription a été effectuée à l'aide de la trousse RiboMAX™ Large Scale RNA Production System, T3 (Promega, Madison, WI). La réaction de synthèse a été réalisée en présence de 4 µl de tampon 5X (Tris-HCl 200 mM, pH 7.5, NaCl 50 mM, MgCl₂ 30mM, spermidine 10 mM), 1.5 µl de 25mM de chaque ribonucléotide (rUTP, rATP, rGTP, rCTP), 3 µg d'ADN linéarisé avec l'enzyme *Bam*H I, 2 µl de 250 µCi/ml [α -³²P-rGTP] (Amersham Ltd., Oakville, Canada) et 2 µl d'enzyme (20 U/µl d'ARN polymérase T3, 15 U/µl de RNasin, 190 U/ml de pyrophosphatase de levure) dans un volume final de 20 µl. La synthèse s'est effectuée à 37°C durant 4 heures. Suite à cette incubation, 2 unités de désoxyribonucléase (RQ1 DNase, Proméga) ont été ajoutés à la réaction. Ce mélange réactionnel a été incubé de nouveau à 37°C durant 30 min. Une extraction phénol/chloroforme (v/v) suivit d'une précipitation à l'isopropanol (0.7v) de la phase aqueuse et de lavages à l'éthanol 70% ont été réalisées suite au traitement à la DNase. L'ARN a été ensuite resuspendu dans 20 µl d'eau contenant 30 unités d'inhibiteur de ribonucléase.

Un volume égal de colorant à la formamide 97% (10 mM EDTA, 1 mg/ml xylène cyanol (xc), 1 mg/ml bleu de bromophénol (bp) dans de la formamide) a été ensuite ajouté. L'échantillon d'ARN a été dénaturé à 65°C 2 min puis placé sur glace aussi pour une durée de 2 min avant d'être déposé sur un gel d'électrophorèse dénaturant (voir Annexes, section 7.1). Après l'électrophorèse, le gel a été exposé à un film Kodak environ 3 min. La bande d'ARN a été excisée du gel et réduite en morceaux à l'aide d'une tige de verre enduite de silicone. L'ARN a été élué avec 400 µl d'une solution 0.5 M acétate d'ammonium et 0.1% SDS toute la nuit à 4°C. Le surnageant a été passé sur une colonne Séphadex G-50, l'ARN précipité à l'isopropanol (0.7v), lavé à l'éthanol 70% et conservé sous forme de culot lyophilisé.

2.5.6 Synthèse *in vitro* de la sonde d'ADN GAPDH

Le plasmide utilisé pour la synthèse *in vitro* de la sonde d'ADN GAPDH a été généreusement fourni par le Dr Jean Morisset. Cette construction correspond au plasmide pUC19 (New England BioLabs Inc.) dont l'insert est l'ADNc correspondant à une partie de l'ARNm de l'enzyme humaine glycéraldéhyde-3-phosphate-déshydrogénase (GAPDH) cloné aux sites de restriction *Xba* I et *Hind* III.

La synthèse de la sonde d'ADN GAPDH a été réalisée par extension d'amorces. Le plasmide a dû subir, au préalable, une digestion par les enzymes *Xba* I et *Hind* III, qui flanquent l'insert, afin que l'ADN généré lors de l'extension d'amorces corresponde uniquement à l'insert cloné. La réaction a été réalisée à l'aide de 100 ng d'ADN correspondant à l'insert digéré et 5 µg d'oligonucléotides dégénérés (pd(N)6) (Amersham, Pharmacia Biotech Inc.), dans un volume final de 15 µl. Ce mélange réactionnel a été incubé 2 min à 100°C. Suite à une incubation de 5 min sur la glace, 2.5 µl d'un mélange des nucléotides dCTP, dGTP et dTTP, chacun à une concentration de 25mM, a été ajouté au mélange réactionnel. La réaction a eu lieu dans un tampon OPA 1X (10 mM Tris-acétate, pH 7.5, 10 mM Magnésium acétate, 50 mM Potassium acétate) en présence de 20 µCi [α -³²P-dATP] (Amersham Ltd., Oakville, Canada) à l'aide de 6 unités de l'ADN polymérase Klenow (Amersham Pharmacia Biotech Inc.). Suite à une incubation de 2 heures à 37°C, les nucléotides libres ont été séparés de la sonde d'ADN radiomarquée sur une colonne Séphadex G-50.

2.5.7 Hybridation des sondes

La préhybridation et l'hybridation des sondes sur les membranes respectives ont été réalisées dans un tampon contenant le réactif Denhardt's 5X, SSC 6X, SDS 0.5% et 100 µg/ml d'ADN de sperme de saumon préalablement chauffé 10 min à 100°C et refroidit sur glace 5 min. La préhybridation et l'hybridation ont été faites à 68°C durant 6 et 16 heures respectivement. La sonde d'ARN AgHD-S a été incubée 1 min à 90°C suivit de 10 min sur la glace. La sonde d'ADN GAPDH a été incubée 10 min à 90°C suivit de 10 min sur la glace. Un million de comptes par minute (cpm) de sonde radiomarquée ont été ajoutés à la solution d'hybridation.

Suite à l'hybridation, différentes étapes de lavages des membranes ont été réalisées: 1) 1 lavage avec du SSC 2X, 10 min à la température de la pièce, 2) 2 lavages avec une solution de SSC 1X et de SDS 1%, 20 min à 55°C, 3) 1 lavage avec une solution de SSC 0.5X et de SDS 0.5%, 20 min à 55°C, 4) 1 lavage avec une solution de SSC 0.1X et de SDS 0.1%, 15 min à la température de la pièce. Après les étapes de lavage des membranes, celles-ci ont été placées dans un sac et exposées toute la nuit à 4°C dans une cassette permettant la révélation par «phosphorimager» (Molecular Dynamic).

3.0 RÉSULTATS

3.1 COMPILATION DE SÉQUENCES

Le laboratoire a conçu une compilation de séquences adaptée aux besoins de la communauté scientifique. Cette banque contient toutes les séquences connues des viroïdes ainsi que celles d'autres ARNs apparentés publiées dans les journaux scientifiques ou disponibles dans GenBank ou dans la librairie de séquences nucléotidiques EMBL. Une section de ce répertoire de séquences est réservée au VHD. Le but de la conception de cette banque est de faciliter les études phylogéniques et de structure-fonction concernant les viroïdes et autres ARNs apparentés comme l'ARN viral du VHD. Différentes informations relatives à ces espèces d'ARN sont retrouvées dans la banque. Par exemple, on retrouve la position des motifs autocatalytiques, des informations sur la structure secondaire ainsi que des renseignements concernant la phylogénie de ces espèces d'ARN. Cette banque est disponible par Internet à l'adresse suivante: <http://www.callisto.si.usherb.ca/~jpperra> (Lafontaine et al., 1999). L'analyse et la description de celle-ci sont décrites sous forme de publications littéraires (voir en Annexes, sections 8.4 et 8.5) (Lafontaine et al., 1998, 1999). Ces publications traitent de la nouvelle nomenclature proposée ainsi que de certains aspects relatifs à la phylogénie et à la structure de ces ARNs.

La section réservée au virus de l'hépatite *delta* humaine, dont j'ai participé à la création, contient toutes les séquences virales connues de ce virus (vHDV).

Contrairement à la section portant sur les viroïdes, on y retrouve les séquences complètes ainsi que les séquences partielles connues du VHD (Figure 7A). Dans la majorité des cas, les séquences partielles correspondent soit aux domaines autocatalytiques ou au cadre de lecture ouvert retrouvé dans le génome du VHD. Actuellement, 23 séquences complètes et 165 séquences partielles du VHD ont été identifiées et répertoriées. La figure 7B est un exemple de l'information retrouvée dans la section attribuée au VHD. Les renseignements portent, entre autre, sur les caractéristiques structurales du génome. On y retrouve des renseignements concernant la polarité de la séquence virale et la position des sites autocatalytiques. Il y a aussi de l'information relative à l'antigène *delta* comme la position des codons d'initiation et de terminaison de la transcription ainsi que le site d'édition. Enfin, la séquence nucléotidique et la provenance de celle-ci sont présentées.

A)

RNA species	Number of sequences
Complete sequences..... part 1 , part 2	23
Partial sequences..... part 1 , part2 ., part3 , part4	165
View a sequence alignment.	
View sequence variation in <i>delta</i> self-cleaving motifs.	

B)

vHDV.5

accession numbers: M21012(gb), 329989(gi)

Locus HPDGEN

1679 nucleotides (332A, 514C, 493G, 340T)

Kuo, M.Y.P., Goldberg, J., Coates, L., Mason, W.S., Gerin, J. and Taylor, J. Molecular cloning of hepatitis *delta* virus RNA from an infected woodchuck liver: Sequence, structure, and applications. *J. Virol.* **62**, 1855-1861 (1988)

```

CCTGAGCCAA GTTCCGAGCG AGGAGACGCG GGGGGAGGAT CAGCACCCGA GAGGGGATGT CACGGTAAAG
AGCATTGGAA CGTCGGAGAA ACTACTCCCA AGAAGCAAAG AGAGGTCTTA GGAAGCGGAC GAGATCCCCA
CAACGCCGGA GAATCTCTGG AAGGGGAAAG AGGAAGGTGG AAGAAAAAGG GGCGGGCCTC CCGATCCGAG
GGGCCCAATC TCCAGATCTG GAGAGCACTC CGGCCCGAAG GGTTGAGTAG CACTCAGAGG GAGGAATCCA
CTCGGAGATG AGCAGAGAAA TCACCTCCAG AGGACCCCTT CAGCGAACAA GAGGCGCTTC GAGCGGTAGG
AGTAAGACCA TAGCGATAGG AGGAGATGCT AGGAGTAGGG GGAGACCGAA GCGAGGAGGA AAGCAAAGAA
AGCAACGGGG CTAGCCGGTG GGTGTTCCGC CCCCCGAGAG GGGACGAGTG AGGCTTATCC CGGGGAATC
GACTTATCGT CCCCATCTAG CGGGACCCCG GACCCCTTTC GAAAGTGACC GGAGGGGGTG CTGGGAACAC
CGGGGACCAG TGGAGCCATG GGATGCCCTT CCGATGCTC GATTCCGACT CCCCCCCCCA AGGGTCGCCC
AGGAATGGCG GGACCCCACT CTGCAGGGTC CGCGTTCCAT CCTTTCTTAC CTGATGGCCG GCATGGTCCC
AGCCTCCTCG CTGGCGCCGG CTGGGCAACA TTCCGAGGGG ACCGTCCCCCT CGGTAATGGC GAATGGGACC
CACAAATCTC TCTAGATTCC GATAGAGAAT CGAGAGAAAA GTGGCTCTCC CTTAGCCATC CGAGTGGACC
TGCGTCTCTC TTCGGATGCC CAGGTCTGGAC CGCGAGGAGG TGGAGATGCC ATGCCGACCC GAAGAGGAAA
GAAGGACGCG AGACGCAAAC CTGTGAGTGG AAACCCGCTT TATTCACTGG GGTGCAACAAC TCTGGGGAGA
AAAGGGCGGA TCGGCTGGGA AGAGTATATC CTATGGAAAT CCCTGGTTTC CCCTGATGTC CAGCCCCTCC
CCGTCCGAG AGAAGGGGGA CTCCGGGACT CCCTGCAGAT TGGGGACGAA GCCGCCCCCG GGCGCTCCCC
TCGATCCACC TTCGAGGGGG TTCACACCCC CAACCGGCGG GCCGGCTACT CTTCTTTCCC TTCTCTCGTC
TTCTCTGGTC AACCTCCTGA GTTCCTCTTC TTCTCTCTTG CTGAGGTTCT TGCCCTCCGC CGATAGCTGC
TTCTTCTTGT TCTCGAGGGC CTTCCTTCGT CGGTGATCCT GCCTCTCCTT GTCGGTGAAT CCTCCCCTGA
GAGGCCTCTT CCGAGGTCCG GAGTCTACCT CCATCTGGTC CGTTCGGGCC CTCTTCGCGG GGGGAGCCCC
CTCTCCATCC TTATCCTTCT TTCCGAGAAT TCCTTTGATG TTCCCCAGCC AGGGATTTTC GTCTCTATC
TTCTTGAGTT TCTCTTTTGT CTCCGGAGG TCTCTCTCGA GTTCCTCTAA CTTCTTTCTT CCGGCCACCC
ACTGCTCGAG GATCTCTTCT CTCCCTCCGC GGTTCCTCCT CGACTCGGAC CGGCTCATCT CGGCTAGAGG
CGGCAGTCCT CAGTACTCTT AACTTTTCT GTAAAGAGGA GACTGCTGGA CTCGCCGCC GAGCCCCGAG

```

Structural characteristics

Genome genomic polarity complete sequence	HDV antigen start codon position : 1598 stop codon position : 954 predicted RNA-edition site : 1012 RNA-edition codon : codon stop
(+) polarity-> Delta ribozyme structural domain (686-769) cleavage site (685-686)	(-) polarity-> Delta ribozyme structural domain (817-900) cleavage site (900-901)

Figure 7: Compilation des séquences et informations relatives au VHD. A) La compilation des séquences complètes et partielles connues ainsi que l'alignement de séquences correspondant à certains domaines du VHD sont disponibles à l'adresse Internet suivante: <http://www.callisto.si.usherb.ca/~jpperra> ou sous forme de publications littéraires. B) Exemple d'une séquence complète (vHDV.5) et de l'information relative au VHD retrouvée dans la banque.

3.2 ALIGNEMENT ET ANALYSE DES SÉQUENCES COMPLÈTES DU VHD

L'alignement précis des séquences complètes connues du VHD a été réalisé. Cet alignement de séquences démontré en annexe à la section 7.1, ne fait pas la distinction entre les différents génotypes. Or, des analyses phylogénétiques préliminaires semblent indiquer que les séquences 9, 20 et 22 appartiennent au génotype II, que la séquence 14 appartient au génotype III et que les 18 autres séquences sont de génotype I. L'analyse des séquences alignées révèle la présence de trois domaines au sein du génome viral du VHD, dans les trois génotypes, se distinguant par leur degré de variabilité. La figure 8 est une représentation schématique des trois domaines retrouvés au sein du génome viral du VHD. Le premier domaine allant des nucléotides 1600 à 1672 et 1 à 655 est très variable. Lorsqu'on analyse le génome en entier, les isolats appartenant au génotypes II et III présentent respectivement 80% et 65% d'homologie avec ceux appartenant au génotype I. Cette variabilité entre les séquences appartenant aux trois différents génotypes, aussi nommée hétérogénéité, est particulièrement observable dans le domaine variable. Jusqu'à ce jour, aucune fonction biologique n'a été attribuée à cette région du génome du VHD ce qui explique possiblement la présence de nombreuses variations nucléotidiques entre les séquences des trois génotypes. Le deuxième domaine allant de la position 656 à 948 est le domaine «viroid-like», nommé ainsi en raison de sa ressemblance avec les viroïdes. La séquence correspondant à ce domaine est relativement bien conservée mais présente quelques variations entre les trois génotypes. En fait, les isolats appartenant aux génotypes II et III présentent respectivement 95% et 85% d'homologie avec ceux du

génotype I. La variation de séquence nucléotidique dans le domaine «viroid-like» est moins importante que celle dans le domaine variable possiblement à cause du rôle essentiel des motifs autocatalytiques dans la réplication du VHD. Le troisième domaine possède le cadre de lecture ouvert encodant les deux isoformes de l'antigène *delta* et comprend les nucléotides 949 à 1599 (Figure 8). Il s'agit de deux protéines essentielles à la réplication et à l'encapsidation du VHD. Tout en étant très conservée, la séquence de ce domaine est légèrement plus variable que celle du domaine «viroid-like» en ce qui concerne les séquences appartenant au génotype II. En fait, les isolats appartenant aux génotypes II et III présentent respectivement 88% et 85% d'homologie avec ceux du génotype I. La variation de séquence observée dans le domaine «viroid-like» et le domaine codant pourrait possiblement être un facteur permettant d'expliquer les différences de symptômes observés lors d'une infection par le VHD.

Il a été proposé que la courte région allant de la position 1608 à 1669, correspondrait au promoteur pour la réplication de l'ARN du VHD (Beard et al., 1996). Les analyses de cette courte séquence indiquent que les isolats appartenant aux génotypes II et III présentent respectivement 79% et 70% d'homologie avec ceux du génotype I. Dans cette région, la structure secondaire est très bien conservée malgré une variabilité au niveau de la séquence nucléotidique. Il a été admis que la structure est essentielle pour l'activité du promoteur. La section relative au VHD dans la banque de séquences du laboratoire contient aussi un alignement de séquences d'acides aminés relatives au domaine codant. Des études antérieures suggèrent que le taux de substitutions non synonymes est supérieur à celui des substitutions synonymes (Krushkal et Li, 1995). La

préférence des nucléotides G et C à la troisième position dans un codon pourrait expliquer la diminution du taux de substitutions synonymes. La compréhension de la nature des forces de sélection impliquées permettrait possiblement de mieux comprendre la biologie du VHD. Enfin, le patron des nucléotides à l'intérieur du génome d'ARN du VHD a été investigué par un groupe de recherche. Les résultats préliminaires ont établi que le génome du VHD est composé d'un grand nombre de segments caractérisés par des tracés de polypurines et de polypyrimidines (Branch et al, 1993). L'analyse du patron des nucléotides effectuée à l'aide de notre alignement de séquences n'indique aucune différence entre les trois génotypes. La signification biologique de la présence de ces signatures au sein du génome du VHD demeure toujours inconnue.

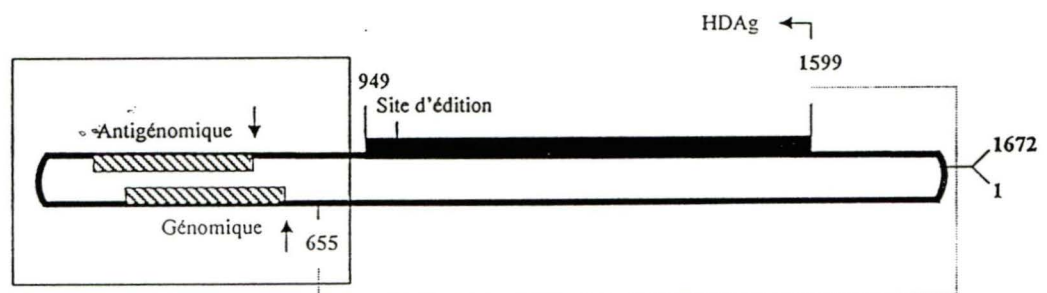


Figure 8: Représentation schématique des trois domaines retrouvés au sein du génome viral du VHD. Le domaine de droite allant des nucléotides 1600 à 1672 et 1 à 655 est variable. Le domaine terminal de gauche qui est encadré allant de la position 656 à 948 est le domaine « viroid-like ». Le troisième domaine représenté par le rectangle noir constitue le cadre de lecture ouvert encodant les deux isoformes de l'antigène *delta*. La séquence de ce domaine comprend les nucléotides 949 à 1599. La numérotation correspond à celle du variant vHDV.5 (Lafontaine et al., 1997).

3.3 SÉQUENÇAGE D'UN ISOLAT CANADIEN DU VHD

3.3.1 Stratégie de clonage

L'établissement d'une relation entre les génotypes, le degré de variabilité des séquences, la distribution géographique et le niveau de sévérité de la maladie nécessite un grand nombre de séquences virales du VHD. De façon à contribuer à cette étude, la séquence nucléotidique complète du VHD isolé d'un sujet canadien infecté par ce virus a été déterminée. La synthèse d'un ADNc à l'ARN viral extrait suivit d'une amplification par PCR furent réalisées. La stratégie adoptée impliquait l'utilisation de quatre paires d'amorces différentes. Comme le montre la figure 9, cette stratégie permet la synthèse de quatre fragments d'amplification de longueur variable allant de 125 à 838 paires de bases. Le chevauchement de ces fragments permet d'amplifier la totalité du génome viral. L'approche impliquant l'amplification par PCR de quatre fragments facilite le séquençage subséquent et constitue une première façon de diminuer la probabilité d'artéfacts dérivés du PCR puisque les fragments obtenus sont moins longs. La deuxième façon était d'utiliser une enzyme polymérase *Taq* purifiée au laboratoire et d'effectuer un nombre maximal de 30 cycles d'amplification.

La transcription inverse a nécessité l'essai des amorces antisens de chacune des paires d'amorces. Celles-ci ainsi que plusieurs autres furent essayées sans succès. Or, malgré la forte structure secondaire du génome viral, une amorce permettant la synthèse d'un ADNc correspondant à la séquence complète du VHD fut identifiée (AS1100). Cette

amorce s'hybride à l'ARN viral entre les nucléotides 1163 et 1180. L'amplification PCR subséquente s'est faite à l'aide des quatre paires d'amorces. Comme l'amorce AS1100, les amorces utilisées pour les réactions de PCR ont été déterminées à l'aide de l'alignement des séquences complètes du VHD dans des régions du génome viral où la séquence est hautement conservée pour les trois génotypes. La séquence de celles-ci ainsi que leur position respective au sein du génome viral sont montrées dans le tableau 2.

Tableau 2: Séquence et position des amorces de RT-PCR et de séquençage.

Nom	Séquence	Position
Amorces sens		
S200	5'dCGATCCGAGGGGCCCAAC3'	202-219
S275	5'dCACTCCGGCCCGAAGGGTTG3'	236-255
S700	5'dTGGCCGGCATGGTCCCAGCC3'	685-704
S1100	5'dGCCGCCCCCGGGCGCTCC3'	1101-1118
Amorces antisens		
AS275	5'dCAACCCTTCGGGCCGGAGTG3'	255-236
AS300	5'dGCTGAAGGGGTCCTCTGG3'	324-304
AS800	5'dCGCACGTCCACTCGGATG3'	844-827
AS1100	5'dGGAAAAGAAGAGTAGCCG3'	1180-1163
AS1200	5'dGAAGGAAGGCCCTCGAGAACAAG3'	1288-1266
Amorces de séquençage		
F606	5'dGAAGGCGGATCGAGGGGAGC3'	1133-1114
F606N	5'dCTTTCCTCTTCGGGTCGG3'	911-894
F836N	5'dGCTTCTTGGGAGTAGTTTC3'	106-88
F836Z	5'dGAGGACGAAAATCCCTGG3'	1466-1449

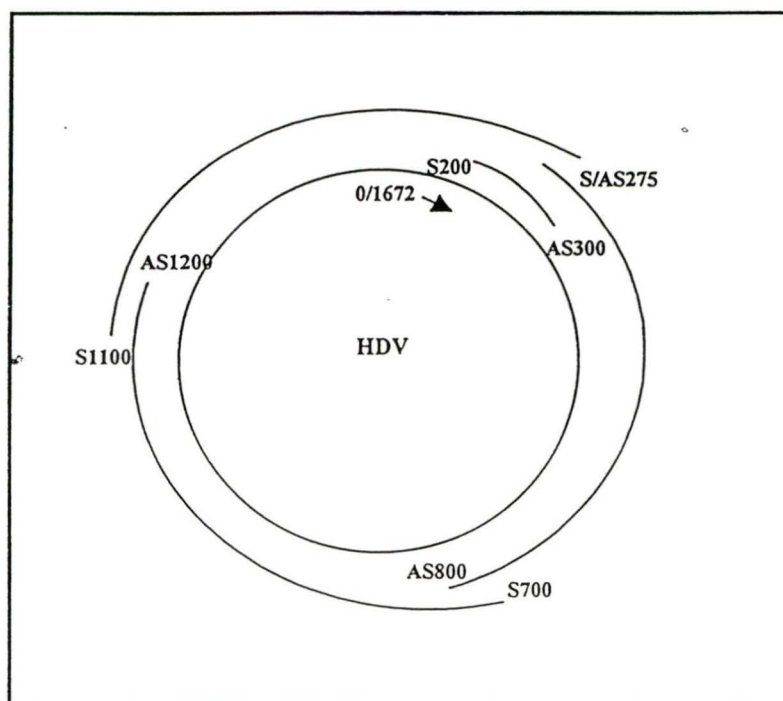


Figure 9: Représentation schématique des fragments amplifiés par PCR. Le cercle représente l'ADNc de l'ARN viral circulaire du VHD. Les quatre fragments amplifiés par PCR sont schématisés par les courbes à l'extérieur de l'ADNc. Les positions des amorces servant pour la synthèse des quatre fragments par PCR sont identifiées aux extrémités de ceux-ci. La flèche représente l'orientation de l'ARN de polarité génomique.

Comme le montre la figure 10, les quatre produits d'amplification attendus ont été obtenus suite à une séparation par électrophorèse sur gel d'agarose. De plus, l'intensité des bandes séparées sur le gel est proportionnelle à la longueur du produit d'amplification obtenu par PCR. Par exemple, le produit d'amplification de 838 paires de bases (piste 8) est plus intense que celui de 125 (piste 12). Des contrôles positifs réalisés à l'aide d'un dimère du VHD (pHDV (D2)) et des quatre paires d'amorces ont servi d'indication pour déterminer les produits de PCR obtenus. Le plasmide utilisé pour obtenir les contrôles positifs est un sous-clone de la construction pSVL (D3) qui a été généreusement fourni par le Dr John Taylor. (Kuo et al.,1989). Une série de contrôles négatifs ont été faits de façon à éviter une mauvaise analyse des résultats. Tous ces contrôles furent réalisés à l'aide des quatre paires d'amorces mais seul ceux faits avec la paire d'amorces S700/AS1200 sont montrés sur le gel. Le premier contrôle négatif est un RT-PCR sans ARN. Le deuxième est une transcription inverse effectuée sans amorce antisens. Cette amorce a été utilisée seulement dans la réaction de PCR. Le troisième contrôle négatif est un RT-PCR dans lequel la réaction de transcription inverse a été faite avec l'amorce sens. Enfin, le quatrième contrôle négatif est un PCR réalisé sans faire de RT au préalable. Les quatre paires d'amorces utilisées pour la synthèse et l'amplification de l'ADNc ont permis d'obtenir quatre fragments. Comme ces produits de RT-PCR se chevauchent, il a été possible d'amplifier entièrement le génome viral du VHD. La séquence des amorces correspond à des régions très conservées dans les trois génotypes du VHD. Les résultats obtenus suggèrent que les amorces utilisées dans cette étude seraient efficaces pour l'amplification de n'importe quelle séquence virale du VHD.

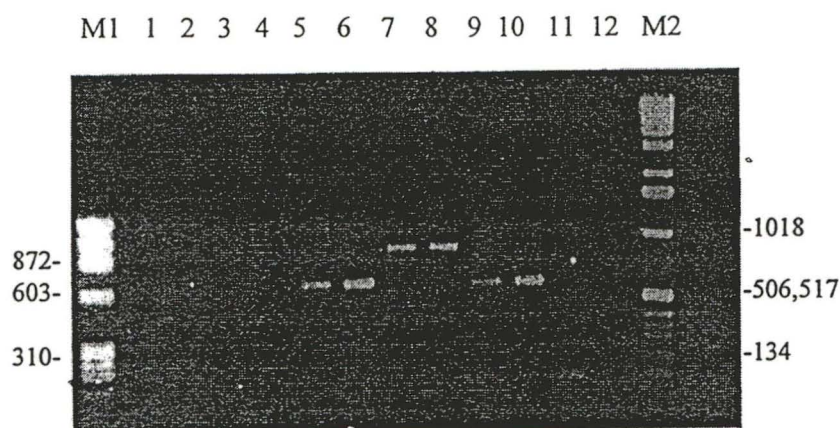


Figure 10: Fragments d'amplification obtenus par RT-PCR. Les quatre produits d'amplification obtenus par RT-PCR ont été séparés par électrophorèse sur gel d'agarose 1.5% et visualisés par coloration au bromure d'éthidium. Les pistes 1 à 4 sont des contrôles négatifs réalisés avec la paire d'amorces S700/AS1200: piste 1, RT-PCR sans ARN, piste 2, réaction RT sans amorce antisens. Cette amorce a été utilisée seulement dans la réaction de PCR, piste 3, RT-PCR dans lequel la réaction de transcription inverse a été faite avec l'amorce sens, piste 4, PCR sans faire de RT. Des contrôles similaires ont été réalisés avec les autres paires d'amorces PCR. Ceux-ci n'ont pas été inclus dans ce gel. Les pistes 5, 7, 9 et 11 sont des contrôles positifs réalisés à l'aide d'un dimère du VHD (pHDV (D2)). Les pistes 6, 8, 10 et 12 sont les produits de RT-PCR obtenus à partir de l'ARN viral du VHD extrait du sérum d'un sujet canadien infecté. Les paires d'amorces utilisées pour les réactions de PCR sont: pistes 5 et 6, S700/AS1200 (606 pb); pistes 7 et 8, S1100/AS275 (838 pb); pistes 9 et 10, S275/AS800 (610 pb); pistes 11 et 12, S200/AS300 (125 pb). Les pistes M₁ et M₂ sont les marqueurs de poids moléculaire d'ADN ϕ X174 digéré par *Hae* III (Promega) et le 1 kb ladder (Gibco BRL), respectivement.

3.3.2 Clonage et séquençage des produits d'amplification RT-PCR

Dans le but de déterminer la séquence nucléotidique complète du VHD isolé d'un sujet canadien, les quatre produits obtenus par RT-PCR ont été clonés et séquencés séparément. En raison de la longueur relativement grande de certains inserts, le séquençage complet de ceux-ci a nécessité l'utilisation d'amorces de séquençage (voir Tableau2, section 3.3.1 des Résultats). Comme les amorces de RT-PCR, les amorces de séquençage furent déterminées à l'aide de l'alignement des séquences complètes du VHD. La séquence et la position de ces amorces sont montrées dans le tableau 2. La séquence nucléotidique complète de cet isolat canadien est représentée dans la figure 11. Cette nouvelle séquence du VHD, nommée vHDV.23, a été ajoutée à la banque de données du laboratoire à l'adresse Internet suivante: <http://www.callisto.si.usherb.ca/~jpperra> (Lafontaine et al. 1999) et à l'alignement des séquences du VHD (voir en annexes, section 7.2). Elle est aussi disponible dans Genbank; dans la banque du «National Center for Biotechnology Information» (NCBI) à l'adresse Internet suivante: <http://www.ncbi.nlm.nih.gov/>. Le numéro d'accèsion de cette séquence est: AF098261.

¹CCTGAGCCAAGTTCCGAGCGAGGAGACGCGGGGGAGGATCAGCTCCCGAGAGGGGATGT
⁶¹CACGGTAAAGAGCATTGGAACGTCCGAGAACTACTCCCAAGAAGCAAAGAGAGGTCTTA
¹²¹GGAAGCGGACGAGATCCCCACAACGCCGAGAATCTCTGGAAGGGGAAAGAGGAAGGTGG
¹⁸¹AAGAAAAAGGGGCGGGCCTCCCGATCCGAGGGGCCCAACTAATCTCCAGATCTGGAGAG
²⁴¹CACTCCGGCCCCAAGGGTTGAGAGTACCCAGAGGGAGGAAGCCACACGGAGTAGAACAG
³⁰¹AAAAATCACCTCCAGAGGACCCCTTCAGCGAACAGAGAGCGCATCGCAGAGGGAGTAGAC
³⁶¹CATAGCGATAGGAGGGGATGCTAGGAGTTAGGGGAGACCAAGCGAGGAGAAAGTAAAGAA
⁴²¹GCAACGGGGCTAGCAGGTGGGTGTTCCGCCTCCCGAGAGACGAGTGAGGCTTATCCCGGG
⁴⁸¹GAACTCGACTTATCGTCCCCACATAGCAGGGACCCCGGACCCCTTCGAAAGGTACCGGA
⁵⁴¹GGGGGTGACTCGAACACCGGGGACCAGTGGAGCCATGGGATGCTCTCCCGATGCTCGA
⁶⁰¹TTCCGACTCCCCCAAGGGTCGCCAGGAATGGCGGGACCCCACTCTGCAGGGTCCGCG
⁶⁶¹TTCCATCCTTTCTTACCTGATGGCGGCATGGTCCCAGCCTCCTCGTGGCGCGGCTGGG
⁷²¹CAACATTCCGAGGGGACCGTCCCTCGGTAATGGCGAATGGGACCCACAAACCTCTCTAG
⁷⁸¹CTTCCCAGAGAGAAGCGAGAGGAAAGTGGCTCTCCCTTAGCCATCCGAGTGGACGTGCGT
⁸⁴¹CCTTCTTCGGATGCCAGGTCCGACCGCGAGGAGGTGGAGATGCCATGCCGACCCGAAGA
⁹⁰¹GGAAAGAAGGACGCGAGACGCAAACCTGTGAGTGGAAACCCGCTTTATTCACTGGGGTCG
⁹⁶¹ACAACCTCTGGGGAGAGAAGGGAGGGTGGCTGGGAAGAGTATATCCTATGGGAATCCCTG
¹⁰²¹GCTTCCCCTGATGTCCAGTCCCTCCCGGTCCGAGAGAAGGGGGACTCCGGGACCCCTTGC
¹⁰⁸¹ATGCTGGGGACGAAGCCGCCCGGGCGCTCCCTCGATCCGCCTTCGAGGGGGTTCACA
¹¹⁴¹CCCCAACTGCGGGCCGGCTACTCTTCTTCCCTTCTCTCGTCTTCTCGGTCAACCCCTT
¹²⁰¹TAAGTTCCTCTTCTTCTCCTTGCTGAGGCTCTTCCCCCGCCGATAGTGCTTCTTCT
¹²⁶¹TGTTCTCGAGGGCCTTCTTCTCGTGGTGATCCTGCCTCTCCTTGTGCGTGAATCCTCCCC
¹³²¹TGAGAGGCCTCTTCCAGGTCCGAGTCTACCTCCATCTGGTCCGTTCCGGGCCCTCTTCG
¹³⁸¹CCGGGGGAGCCCCCTCTCCATCCTTATCCTTCTTCCGAGAATTCCTTGATGTTCCCCA
¹⁴⁴¹GCCAGGGATTTTCGTCTCTATCTTCTTGTAGTTTCTTCTTGTCTTCCGAGGTCTCTCT
¹⁵⁰¹CGAGTTCCTCTAACTTCTTCTTCCGGCCACCCACTGCTCGAGGATCTTCTCTCCCTC
¹⁵⁶¹CGCGGTTCTTCTCGACTCGGACCGGCTCATCTCGGCTAGAGGCGGCGAGTCTCAGTACT
¹⁶²¹CTTACTCTTTTCTGTAAAGAGGAGACTGCTGGACTCGCCGCCCGAGCCCGAG

Figure 11: Séquence d'un isolat canadien du VHD. Le site correspondant au codon d'initiation de la transcription et l'orientation de l'AgHD sont indiqués par un trait et une flèche. Les lignes verticales indiquent les sites de terminaison de la transcription pour la forme petite (AgHD-S) et la forme large (AgHD-L) de l'antigène *delta*. Les flèches aux positions 681 et 895 indiquent les sites d'autocoupure des brins génomique et antigénomique respectivement. Les variations nucléotidiques observées (microhétérogénéité) sont indiquées entre parenthèses. Les nucléotides soulignés correspondent à ceux observés pour la première fois. La portion de séquence encadrée et située entre les domaines autocatalytiques correspond à la petite région qualifiée d'hypervariable.

3.3.3 Analyse de la séquence virale de l'isolat canadien du VHD

L'analyse de la séquence révèle qu'elle est composée de 1672 nucléotides. Elle possède un cadre de lecture ouvert permettant d'encoder une protéine fonctionnelle, l'antigène *delta*. De plus, le codon de terminaison UAG est présent à l'endroit approprié dans la séquence de l'AgHD-S. Ce codon peut certainement subir un événement d'édition dans un contexte cellulaire permettant ainsi la transcription et la traduction de la forme large de l'antigène *delta* (AgHD-L). Cette séquence a été alignée aux 22 autres séquences complètes connues du VHD (voir en Annexes, section 7.2). La comparaison de la séquence de l'isolat canadien avec les autres séquences permet d'observer de la variabilité à travers tout le génome. De plus, cette séquence est typique du génotype I. En effet, la séquence possède plus de 80% d'homologie avec les autres séquences appartenant au génotype I. De plus, l'homologie de séquence avec les isolats du génotype I est similaire pour les trois domaines retrouvés dans le génome du VHD. Cette variabilité confirme que l'isolat canadien de même que les autres séquences isolées de divers régions sont hétérogènes.

Dans le but de vérifier la présence de microhétérogénéité, c'est-à-dire, de la variation de séquences virales au sein d'un même sujet infecté, nous avons déterminé la séquence de quatre clones respectifs de chaque fragment cloné. L'analyse de ces séquences et la comparaison entre elles indiquent la présence de microhétérogénéité. En fait, on observe des différences entre chacun des quatre clones. Au total, des variations

ont été identifiées à 17 positions dans le génome viral, correspondant à une variation de 1% (17 variations dans une molécule de 1672 nt). Cette variabilité de séquence était observable dans tout le génome et reflétait les pressions de sélection. Par exemple, la région encodant l'antigène *delta* était très conservée contrairement à la région variable où des variations furent observées. Nous avons identifié 5 mutations observées pour la première fois. Enfin, il est intéressant de noter que 7 des 17 mutations se retrouvent dans la petite région constituée de 48 nucléotides localisée entre les deux motifs autocatalytiques hautement conservés. Cette région est représentée dans un cadre dans la figure 10. Ce résultat nous suggère que cette portion de séquence au sein du génome viral est hypervariable. La *Taq* polymérase effectue 2×10^{-4} mauvaises incorporations de nucléotides par cycle, après le trentième cycle (Saiki et al., 1988). Comme nous avons effectué 30 cycles et que le taux de variation observé est supérieur aux taux d'erreurs de cette enzyme, la variation observée dans la séquence de l'isolat canadien reflète vraisemblablement la variation réelle entre patients et au sein d'un même sujet infecté. Elle n'est certainement pas due à des erreurs générées lors de l'amplification PCR. Ces résultats ont permis de rédiger un article littéraire accepté pour publication (voir en Annexes, section 8.3) (Poisson et al., 1999)

3.4 CARACTÉRISATION ET IDENTIFICATION DES COLONIES CELLULAIRES EXPRIMANT CONSTITUTIVEMENT L'ARN AgHD-S

Dans le but de mettre au point un outil de travail adéquat pour l'étude des caractéristiques biochimiques du VHD et pour les travaux visant l'inhibition de la

propagation de ce pathogène humain, nous avons établi une nouvelle lignée cellulaire hépatique (HepG2). La particularité de cette lignée réside dans le fait qu'elle exprime constitutivement l'ARNm correspondant à la forme petite de l'antigène *delta* (AgHD-S). L'ARN génomique du VHD produit un unique ARNm qui encode pour les deux isoformes de l'antigène *delta*. Ces deux formes d'antigènes sont indispensables dans le cycle de vie du VHD. En effet, la forme petite (AgHD-S) et la forme large (AgHD-L) de l'antigène *delta* sont impliquées dans la réplication et l'encapsidation du VHD respectivement. Par conséquent, l'inactivation de l'expression de l'ARNm, par exemple, par des ribozymes constituerait une façon efficace de contrôler la propagation du VHD. De plus, l'ARNm AgHD possède moins de structures secondaires que d'autres régions du génome viral comme par exemple les domaines autocatalytiques. Enfin, la séquence de cette portion du génome viral est très bien conservée, ce qui en fait une cible très intéressante pour les études d'inhibition du VHD.

Les motifs autocatalytiques retrouvés au sein du génome et indispensables à la réplication du VHD peuvent être modifiés de façon à les séparer en deux entités: le centre catalytique et le substrat, pour obtenir un système de clivage en *trans*. Cette ingénierie donne naissance au ribozyme *delta* qui peut inactiver successivement plusieurs molécules de substrats en *trans* (revue par Mercure et al., 1997). Puisqu'il est théoriquement possible de construire des ribozymes contre n'importe quel ARN, de le reconnaître spécifiquement par appariement de paires de bases et enfin de l'inactiver par clivage à un site spécifique, les ribozymes ont d'énormes applications potentielles (revue par James et Gibson, 1998). Des travaux antérieurs effectués au laboratoire, dans lesquels l'ARNm du

VHD a été ciblé par un ribozyme *delta in vitro*, ont permis d'identifier des sites de coupure potentiels sur cet ARN. Des ribozymes spécifiques à ces sites de clivage ont été fabriqués et caractérisés au laboratoire. L'utilisation de ces ribozymes dans un contexte cellulaire permettrait d'approfondir nos connaissances sur les facteurs impliqués dans l'activité de coupure *in vivo*. Ainsi, la lignée hépatique établie sera un outil de travail pour le développement des moyens permettant l'inactivation de la propagation de ce pathogène humain.

Cette lignée fut développée par la transfection stable d'un vecteur d'expression préalablement construit au laboratoire (pcDNA3-Ag-S). La recherche des colonies cellulaires exprimant l'ARNm AgHD-S de façon constitutive a été réalisée. Dans un premier temps, une réaction de transcription inverse effectuée sur l'ARN cytoplasmique extrait des colonies cellulaires a été faite à l'aide d'une amorce spécifique à l'AgHD-S. Cette amorce s'hybride à une extrémité de l'ARNm et assure la synthèse d'un ARN complet. Ensuite, une amplification par PCR du produit de la réaction de RT a été réalisée à l'aide de deux amorces spécifiques qui s'hybrident aux deux extrémités de l'ARNm. Ceci a permis d'obtenir un produit d'amplification correspondant à la totalité de l'ARNm. La première série de pistes montre le résultat de la séparation par électrophorèse sur gel d'agarose des produits de PCR obtenus (Figure 12 A). L'identification des colonies positives a été faite à l'aide d'un contrôle positif. En fait, une transcription *in vitro* à partir d'une construction (pKSAgS) dans laquelle l'insert AgHD-S a été cloné a permis la synthèse de l'ARN HDAg-S, lequel a été amplifié par RT-PCR. Deux types de contrôles négatifs ont été faits. Un premier contrôle négatif a été

réalisé à partir d'un extrait d'ARN cytoplasmique provenant d'une colonie cellulaire dans laquelle le vecteur pcDNA3 sans insert a été transfecté de façon stable. Une réaction de RT-PCR faite à partir d'ARN extrait de cellules HepG2 non transfectées constitue le deuxième contrôle négatif. Une réaction de PCR directement à partir d'extraits d'ARN provenant soit des colonies 1 et 2, de la colonie transfectée par le vecteur pcDNA3 seul et des cellules non transfectées permet de s'assurer que l'amplification PCR s'est faite à partir de l'ADNc AgHD-S et non de l'ADN génomique. La deuxième série de piste correspond à un contrôle de l'expression du gène de l'enzyme humaine glycéraldéhyde-3-phosphate-déshydrogénase (GAPDH) (Figure 12 A)). L'analyse des résultats obtenus par RT-PCR révèle qu'il y a deux colonies qui expriment constitutivement l'ARN AgHD-S, la colonie numéro 1 (piste 4 de la première série de pistes) ayant un taux d'expression environ 3 à 5 fois plus élevé que la seconde (piste 6). De façon à confirmer ces résultats, nous avons fait un buvardage de type Southern sur les produits de PCR (Figure 12 B). L'hybridation de la membrane, réalisée à l'aide de sondes spécifiques à l'AgHD-S et au GAPDH, a permis de confirmer que les deux colonies identifiées par RT-PCR sont positives. En d'autres termes, ces colonies expriment l'ARN AgHD-S constitutivement.

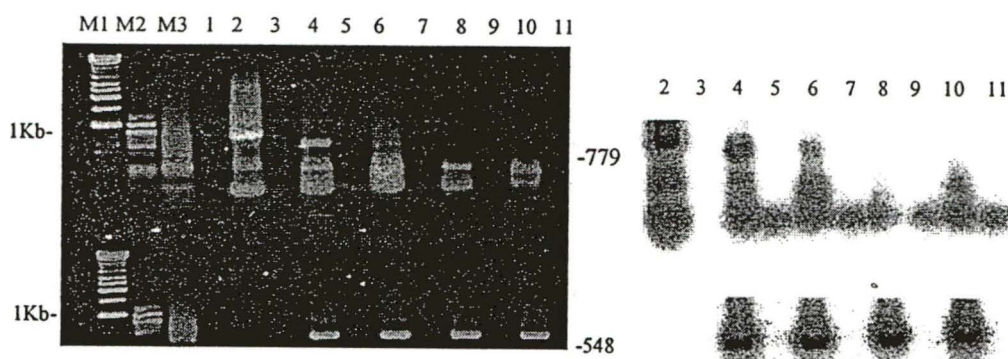


Figure 12: RT-PCR réalisés à partir de l'ARN cytoplasmique extrait des colonies cellulaires. A) Électrophorèse en gel d'agarose des produits de RT-PCR obtenus à partir de l'ARN cytoplasmique extrait des colonies cellulaires HepG2. La première série de pistes correspond aux produits de RT-PCR obtenus à l'aide des amorces spécifiques à l'AgHD-S et la deuxième série de piste correspond aux produits obtenus à l'aide d'amorces spécifiques au GAPDH. La piste 2 de la première série de pistes est le contrôle positif fait à l'aide du plasmide pKSAgS. Dans les deux séries de pistes, les puits 4 et 6 sont les produits de RT-PCR fait à partir des colonies 1 et 2 respectivement, dans lesquelles le vecteur pcDNA3-AgHD-S a été transfecté. La piste 8 est un contrôle négatif correspondant à la colonie cellulaire dans laquelle le vecteur pcDNA3 sans insert a été transfecté. La piste 10 est un contrôle négatif dans lequel un RT-PCR a été fait à partir de l'ARN extrait de cellules non transfectées. Les pistes 5, 7, 9, 11 sont des contrôles négatifs dans lesquels un PCR a été fait à partir de l'ARN extrait, soit des colonies 1 et 2, de la colonie ayant été transfectée par le vecteur pcDNA3 seul et des cellules non transfectées. B) Hybridation de la membrane issue du transfert de type Southern des produits de RT-PCR séparés en gel d'agarose.

4.0 DISCUSSION

L'objectif du présent travail était de développer des outils pour faciliter l'étude et la compréhension des propriétés biologiques du virus de l'hépatite *delta* humaine (VHD). Afin de réaliser ceci, trois avenues de recherche furent utilisées. Dans un premier temps, nous avons participé à la création d'une banque de séquences contenant une section spécifique au VHD. La compilation des séquences ainsi que plusieurs éléments structuraux relatifs au VHD ont été mis à la disposition de la communauté scientifique, ces informations étant fréquemment mises à jour (voir Résultats, sections 3.1 et 3.2). De plus, un alignement exhaustif des 22 séquences virales connues de ce virus a été produit et ajouté à la banque. En deuxième lieu, de façon à contribuer à l'étude des mécanismes d'évolution du VHD, nous avons déterminé la séquence nucléotidique complète d'un isolat canadien à l'aide d'une stratégie de clonage impliquant l'utilisation de quatre paires d'amorces en RT-PCR (voir Résultats, section 3.3). Les études relatives à l'évolution du VHD nécessitant un grand nombre de séquences, l'isolat canadien constitue un échantillon de plus pouvant être utilisé. Dans un troisième temps, nous avons établi une nouvelle lignée hépatique qui exprime constitutivement l'ARNm encodant la forme petite de l'antigène *delta* (AgHD-S) (voir Résultats, section 3.4). Cette lignée constitue un outil idéal, par exemple, pour des essais d'inactivation génique dans lesquels des drogues constituées d'acides nucléiques seraient utilisées. Bien qu'à première vue ces trois contributions peuvent apparaître sans relation, elles sont étroitement associées et leur interdépendance est indéniable.

4.1 SÉQUENCE DU VHD

La compilation de séquences ainsi que les informations structurales retrouvées dans la banque constituent un outil de référence pour diverses recherches. Par exemple, une compilation des variations de séquence naturelles relatives aux deux motifs autocatalytiques *delta*, a été ajoutée récemment. Tout dépendant de leur position, certaines bases supportent une variation tandis que d'autres sont toujours conservées. Les bases qui sont hautement conservées contribueraient au mécanisme moléculaire de l'activité de coupure des motifs autocatalytiques. Dans la présente étude, l'alignement précis des séquences complètes du VHD a permis de confirmer la présence de trois domaines qui se distinguent par leur degré de variabilité au sein du génome viral. De plus, il a permis de déterminer les amorces spécifiques au VHD qui ont le potentiel d'amplifier n'importe quelle séquence virale peu importe le génotype. Ces amorces ont permis d'amplifier avec succès un isolat d'un sujet canadien infecté par le VHD (voir Résultats, section 3.3). Le diagnostic d'une infection par le virus de l'hépatite D se fait, entre autre, par la recherche de l'ARN viral via une amplification par RT-PCR. Étant donné le caractère variable de certaines régions du génome du VHD, il est impératif de choisir des amorces qui s'hybrident dans des régions hautement conservées permettant ainsi de détecter l'ARN chez un nombre maximal de porteurs. L'utilisation des amorces déterminées dans la présente étude permettrait une identification facile et fiable du génome viral puisqu'elles s'hybrident dans des régions conservées pour tous les génotypes. La banque a donc contribué à la réalisation du projet de séquençage.

Il est tout aussi vrai de dire que la détermination d'un isolat canadien a aussi contribué à parfaire la compilation de séquences du VHD comprises dans la banque du laboratoire. Celle-ci, désormais composée de 23 séquences complètes, nous a permis de vérifier le concept d'hétérogénéité par l'alignement de notre séquence aux 22 autres répertoriées et de faire une analyse d'homologie. Les résultats de cette étude ont révélé l'appartenance de l'isolat canadien au géotype I, lequel est distribué géographiquement dans plusieurs pays (Shakil et al., 1997). Il a été suggéré que l'ARN viral au sein d'un sujet infecté se retrouve sous forme d'une population d'ARN hétérogène (Lee et al., 1992). C'est pourquoi nous nous sommes intéressés à vérifier la présence de microhétérogénéité de séquences virales chez le sujet infecté. En fait, l'utilisation de plusieurs clones correspondant à chacun des fragments amplifiés, isolés et clonés nous a permis d'observer 17 divergences, représentant une variation de 1%. De ces mutations, 7 se retrouvent dans la petite région de 48 nucléotides située entre les deux motifs autocatalytiques du domaine "viroid-like", représentant un taux de variation de 15%. Aucune fonction n'a été attribuée à cette région et très peu d'études de microhétérogénéité ont été faites jusqu'à maintenant. À notre connaissance, notre recherche montre pour la première fois une région hypervariable au sein du génome viral du VHD. L'implication de ce domaine de même que celle des variations génétiques généralement observées dans la détermination de la sévérité de l'infection par le VHD n'est pas claire. Le niveau élevé de variations dans cette courte région pourrait possiblement avoir des conséquences significatives sur la biologie moléculaire et cellulaire du VHD, quoique cette hypothèse demeure incertaine. L'analyse de la distribution géographique des isolats du VHD de même que les variations de séquences

entre isolats et celles observées dans la population d'ARN viral au sein d'un sujet infecté seront importantes pour la compréhension des aspects pathogéniques et épidémiologiques de ce pathogène. À la lueur de nos résultats, il est désormais évident que le domaine hypervariable identifié est une région à proscrire pour le développement de drogues basées sur les acides nucléiques car l'interaction de ces dernières implique la formation de paires de bases avec le cible d'ARN. De plus, le diagnostic d'une infection par le VHD par la recherche de l'ARN viral via une amplification par RT-PCR serait inefficace dans le cas où les amorces utilisées s'hybrideraient dans la région hypervariable.

Enfin, plusieurs groupes émettent l'hypothèse d'une relation entre les génotypes, le degré de variabilité des séquences, la distribution géographique et le niveau de sévérité de la maladie observée suite à une infection par le VHD. En fait, le phénomène de variabilité des séquences entre les sujets infectés et au sein d'un même individu, pourrait contribuer à la diversité des tableaux cliniques observés. Or, ce genre d'étude nécessite un grand nombre de séquences virales. C'est pourquoi, nous avons déterminé la séquence d'un isolat canadien et réalisé une compilation de séquences. Cette dernière constitue un outil indispensable pour tout chercheur dont l'intérêt porte sur l'étude des propriétés et des caractéristiques biologiques du VHD.

4.2 LIGNÉE CELLULAIRE

Nous avons développé une lignée cellulaire hépatique qui permet l'expression constitutive de l'ARNm AgHD-S. Cet ARN constitue une cible idéale pour contrôler la

propagation du VHD étant donné son rôle essentiel dans la réplication du génome viral. De plus, l'ARNm est moins structuré que le génome viral qui adopte une structure semblable à une tige double-brin. Finalement, la séquence de cette portion du génome est très bien conservée. Donc, le ciblage de l'ARNm AgHD-S apparaît tout à fait justifié pour un éventuel traitement du VHD basé sur l'utilisation d'acides nucléiques. La nouvelle lignée hépatique exprimant constitutivement l'ARNm AgHD-S s'avère donc un outil indispensable. Le développement de la lignée cellulaire a impliqué la détermination d'une séquence d'ARNm AgHD-S pour la construction du vecteur d'expression eucaryote utilisé pour la transfection. Parmi toutes les séquences répertoriées dans la banque, nous avons choisi celle du variant vHDV.5 (selon la numérotation de la banque). Cette dernière appartenant au génotype I, reflèterait assez bien un consensus de toutes les autres séquences. En fait, une analyse phylogénétique réalisée au laboratoire suggère que cette séquence serait relativement ancienne et étroitement liée à l'événement qui a suivi à l'apparition des trois génotypes (J.-P. Perreault, communication personnelle). Par conséquent, le développement d'un outil thérapeutique efficace dirigé contre cet ARNm permettrait de cibler presque la totalité des ARNm AgHD-S. Suite à la transfection du vecteur d'expression de l'ARNm AgHD-S, l'identification des colonies positives a impliqué une transcription inverse effectuée sur l'ARN cytoplasmique suivit d'une amplification par PCR. L'avantage de cette stratégie est qu'elle permet une évaluation semi-quantitative de l'expression de l'ARN d'intérêt. En fait, le RT-PCR nous assure que l'amplification se fait à partir de l'ADNc correspondant théoriquement à chacune des molécules d'ARNm présentes dans la cellule. Une amplification par RT-PCR effectuée à partir d'un plasmide contenant l'insert AgHD-S a permis de s'assurer de la validité de

nos résultats. L'amplification ayant été réalisée avec des amorces déterminées à l'aide de l'alignement de séquences dans des régions où la séquence est conservée. Nous avons tenté sans succès de faire la recherche de l'expression de l'ARNm AgHD-S par buvardage de type Northern. La détection de l'ARN d'intérêt par cette approche s'avère parfois ardue étant-donné les pertes de matériel au cours des différentes étapes expérimentales et le niveau de sensibilité relativement faible de la technique.

Le VHD possède des caractéristiques biochimiques qui en font un virus unique chez les mammifères. Parmi celles-ci, il y a les motifs autocatalytique qui catalysent la coupure en *cis* des brins multimériques générant ainsi les monomères lors de la réplication. Ces motifs peuvent être séparés en deux entités. On obtient ainsi le ribozyme *delta* qui a la capacité de reconnaître un substrat et d'opérer une coupure en *trans*. L'habilité de modifier la séquence de reconnaissance du ribozyme *delta* de façon à cibler et à opérer une coupure d'une molécule d'ARN tel un ARN viral ou un ARNm associé à une maladie humaine, lui confère un potentiel thérapeutique très intéressant. Notre groupe de recherche s'intéresse au potentiel thérapeutique du ribozyme *delta* pour contrôler, entre autre, la propagation du VHD.

Dans la poursuite des travaux du laboratoire dans lesquels l'ARNm AgHD-S est ciblé, il a été impératif de déterminer des régions hautement conservées. Des études préliminaires dans lesquelles l'ARNm de l'AgHD-S a été ciblé par un ribozyme *delta* ont permis de déterminer des sites de coupure potentiels sur l'ARN (Roy et al., 1998). L'utilisation de l'alignement de séquences permet le développement de ribozymes *delta*

ayant le potentiel de cibler plusieurs variants de séquence du VHD. La caractérisation *in vitro* de ribozymes pouvant effectuer une coupure spécifique à ces sites a été faite (Roy et al., 1998). Désormais, la lignée cellulaire que nous avons établie sert d'outil pour poursuivre l'étude et la caractérisation de ces ribozymes dans un contexte cellulaire. Des travaux dans lesquels l'efficacité de ces derniers est vérifiée par la diminution du signal de l'ARNm AgHD-S suite au clivage sont présentement en cours. Il serait aussi intéressant de pouvoir vérifier l'efficacité des ribozymes par la diminution de l'expression de l'antigène *delta*. C'est pourquoi la recherche des colonies exprimant cette protéine fait présentement l'objet de travaux dans le laboratoire. De plus, cette lignée pourrait permettre éventuellement de sélectionner directement des ribozymes fonctionnels *in vivo* dans des conditions physiologiques (d'ions, de pH et de tampon), en présence de protéines cellulaires.

En conclusion, les trois aspects du projet dans lesquels des outils de travail ont été développés permettront la poursuite des travaux de recherche fondamentale et clinique relatifs au virus de l'hépatite *delta* humaine.

5.0 REMERCIEMENTS

Je tiens particulièrement à remercier mon père pour sa confiance, son amour et pour la fierté qu'il a toujours su me prodiguer. Il m'attriste qu'il ne soit pas ici pour partager mon bonheur. Sa présence dans mon cœur au cours de ma maîtrise a été pour moi d'un très grand réconfort. Je remercie aussi ma mère et Maxime, mon frère et Annie, ainsi que Benoit et mes amies Guylaine, Mélanie, Judith, Marie-Claude et Isabelle pour leur soutien et leurs encouragements constants tout au long de mes études.

Je tiens également à remercier le Dr Jean-Pierre Perreault pour m'avoir accepté au sein de son équipe comme étudiante à la maîtrise et pour avoir dirigé mes travaux de recherche. J'aimerais aussi remercier Dr Daniel B. Ménard et Dr Éric Frost pour leur précieuse collaboration sans quoi une partie de mon projet de maîtrise n'aurait pu exister. Enfin, je remercie tous mes collègues du département de Biochimie pour leur support technique et pour leurs réponses à mes nombreuses questions.

6.0 RÉFÉRENCES

- Beard, M. R., MacNaughton, T. B., Gowans, E. J. (1996). Identification and characterization of a Hepatitis delta virus RNA transcriptional promoter. *J. Virol.*, **70**:4986-4995.
- Been, M. D. (1994). *Cis*- and *trans*-acting ribozyme from a human pathogen hepatitis delta virus. *Trend Biochem Sci*, **19**:251-256.
- Branch, A. D., Lee, S. E., Neel, O. D., Robertson, H. D. (1993). Prominent polypurine and polypyrimidine tracts in plant viroids and in RNA of the human hepatitis delta virus. *Nucleic Acids Res.* **21**:3529-3535.
- Careno, V., Bartolomé, J., Madejon, A. (1994). Hepatitis *delta* virus infection: molecular biology and treatment. *Dig. Dis.*, **12**:265-275.
- Choo, Q.L., Kuo, G., Weiner, A.J., Oberly, L.R., Bradley, D.W., Houghton, M. (1989). Isolation of a cDNA clone derivated from a blood-borne nonA, non B viral hepatitis genome. *Science*, **244**:359-362.

- Diener, T.O. (1993). Hepatitis *delta* virus: molecular biology, pathogenesis and clinicals issues. (Hadziyannis, S.J., Taylor, J.M., Bonino. Eds Wiley-Liss, New-York).
p.109-116.
- Engelke, D.R., Krikos, A., Bruck, M.E., Ginsgurd, D. (1990). Purification of *Thermus aquaticus* DNA polymerase expressed in *Escherichia coli*. *Analytical Biochemistry*, **191**:396-400.
- Hadziyannis, S.D.(1997). Hepatitis *delta*. *J. Gastroenterology and Hepatology*, **12**:289-298.
- Janot, C. (1996). Diagnostic biologique des hépatites virales. *Gastroenterology Clin Biol.*, **20**:1099-1123.
- James, H.A., Gibson, I. (1998). The therapeutic potential of ribozymes. *Blood*, **91**:371-382.
- Kuo, M.Y., Chao, M., Taylor, J. (1989). Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of *delta* antigen. *J. Virol.*, **63**:1945-50.
- Krushkal, J., Li, W. H. (1995). Substitution rates in hepatitis delta virus. *J. Mol. Evol.* **41**:721-726

Lai, M.M.C. (1995). The molecular biology of hepatitis *delta* virus. *Annu. Rev. Biochem.*, **64**:259-286.

Lafontaine, D.A., Deschênes, P., Bussière, F., Poisson, V., Perreault, J.P. (1999). The viroid and viroid-like RNA database. *Nucleic Acids Res.*, **1**:186-187.

Lafontaine, D.A., Mercure, S., Perreault, J.P. (1998). identification of a Crohn's disease specific transcript with potential as a diagnostic marker. *Gut*, **42**:878-882.

Lafontaine, D.A., Mercure, S., Poisson, V., Perreault, J.P. (1997). The viroid and viroid-like RNA database. *Nucleic Acids Res.*, **1**:190-191.

Lazinski, D.W., Taylor, J.M. (1995). Recent developments in hepatitis *delta* virus research. *Adv. in Virus Res.*, **43**:187-231.

Lee, C.M., Bih, F.Y., Chao, Y.C., Govindarajan, S., Lai, M.M.C. (1992). Evolution of hepatitis *delta* virus RNA during chronic infection. **188**:265-273.

Makino, S., Chang, M.F., Shieh, C.K., kamahora, T., Vannier, D.M., Govindarajan, S., Lai, M.M.C. (1987). Molecular cloning and sequencing of human hepatitis *delta* virus RNA. *Nature*, **329**:343-346.

- Mercure, S., Lafontaine, D., Roy, G., Perreault, J.P. (1997). Le motif autocatalytique d'ARN du virus delta de l'hépatite humaine. *Médecine/Sciences*, **13**:662-669.
- Muerhoff, A.S., Leary, T.P., Simons, J.N., Pilote-Matias, T.J., Dawson, G.J., Erker, J.C. et al. (1995). Genomic organization of GB virus A and B: two new members of the Flaviviridae associated with GB agent hepatitis. *J. Virol.*, **69**:5621-5630.
- Polish, L.B., Gallagher, M., Fields, H.A., Hadler, S.C. (1993). *Delta* hepatitis: Molecular biology and clinical and epidemiological features. *Clin. Microbiol. Rev.*, **3**:211-229.
- Poisson, F., Roingeard, P., Goudeau, A. (1995). Le virus de l'hépatite *delta*: un mode de réplication bien singulier. *Médecine/Sciences.*, **11**:1379-1387.
- Poisson, V., Ménard, D.B., Frost, É., Perreault, J.P. (1999). A canadian isolate of hepatitis *delta* virus. *Can. J. Gastro.* (sous presse).
- Polson, A.G., Bass, B.L., Casey, J.L. (1996) RNA editing of hepatitis *delta* virus antigenome by dsRNA-adenosine deaminase. *Nature.*, **380**:454-6
- Roy, G., Ananvoranich, S., perreault, J.P. (1999). *Delta* ribozyme has the ability to cleave in trans an mRNA. *Nucleic Acids Res.*, **4**:942-948.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B.,

Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.*, **239**:487-491.

Sanger, F., Nicklen, S., Coulson, A.R. (1977). *Proc. Natl. Acad. Sci. USA*, **74**:5463.

Sambrook, J., Fritsch, E.F., Maniatis, T.(1989). Molecular cloning: A laboratory manual.

Second Edition. C. Noland, Editor. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New-York.

Shakil, A.O., Hadziyannis, S., Hoofnagle, J.H., di Bisceglie, A.M., Gerin, J.L., Casey,

J.L. (1997). Geographic distribution and genetic variability of hepatitis *delta* virus Genotype I. *Virology*, **234**:160-167.

Simons, J.N., Leary, T.P., Dawson, G.J., Pilot-Matias, T.J., Muerhoff, A.S., Schlauder,

G.G. et al. (1995). Isolation of a novel virus-like sequences associated with human hepatitis. *Nature Med.* **1**:564-569.

Trépo, C., Valla, D. (1993). Hépatites virales. Paris: Doin. Coll. Progrès en Hépato-

Gastroentérologie (9):188p.

7.0 ANNEXES

7.1 ÉLECTROPHORÈSE EN GEL D'URÉE-POLYACRYLAMIDE

Le gel dénaturant urée-polyacrylamide est composé d'urée (7M), de polyacrylamide à une concentration variable (acrylamide: N,N méthylène-bis-acrylamide; 19:1) et de tampon TBE 1X (90 mM Tris, 10 mM EDTA pH 8.0, 89 mM acide borique). Pour provoquer la polymérisation du gel, 5 μ l de persulfate d'ammonium (PSA) (10%) et 0.5 μ l de TEMED ont été ajoutés à chaque ml de solution. La dimension des gels varie en fonction des exigences de chaque expérience.

7.2 ALIGNEMENT DES 23 SÉQUENCES COMPLÈTES DU VHD

♦

Accession	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

vHDV.1	G	-	G	C	C	C	C	C	C	C	T	T	G	A	A	G	T	G	C	G	G	-	G	-	G		545
vHDV.2	G	A	-	C	C	C	C	C	C	C	T	T	C	A	A	A	T	T	C	C	-	A	-	G	G		546
vHDV.3	G	-	-	T	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		547
vHDV.4	G	-	-	-	-	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		548
vHDV.5	G	-	G	C	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		549
vHDV.6	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		550
vHDV.7	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		551
vHDV.8	G	-	-	C	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		552
vHDV.9	G	-	G	C	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		553
vHDV.10	G	-	G	C	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		554
vHDV.11	G	A	-	G	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		555
vHDV.12	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		556
vHDV.13	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		557
vHDV.14	G	A	T	C	G	A	C	T	T	C	C	T	T	C	A	A	T	T	C	C	-	A	-	G	G		558
vHDV.15	G	-	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		559
vHDV.16	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		560
vHDV.17	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		561
vHDV.18	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		562
vHDV.19	G	A	-	G	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		563
vHDV.20	G	-	-	G	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		564
vHDV.21	G	-	-	G	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		565
vHDV.22	G	A	-	-	-	-	-	-	-	-	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		566
vHDV.23	G	A	-	G	C	C	C	C	C	C	T	T	T	C	A	A	T	T	C	C	-	A	-	G	G		567

[illegible]

Sequence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523
----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

[illegible]

vHDV.1 :
 vHDV.2 :
 vHDV.3 :
 vHDV.4 :
 vHDV.5 :
 vHDV.6 :
 vHDV.7 :
 vHDV.8 :
 vHDV.9 :
 vHDV.10 :
 vHDV.11 :
 vHDV.12 :
 vHDV.13 :
 vHDV.14 :
 vHDV.15 :
 vHDV.16 :
 vHDV.17 :
 vHDV.18 :
 vHDV.19 :
 vHDV.20 :
 vHDV.21 :
 vHDV.22 :
 vHDV.23 :

: 1046
 : 1048
 : 1049
 : 1046
 : 1046
 : 1044
 : 1043
 : 1046
 : 1048
 : 1045
 : 1046
 : 1047
 : 1048
 : 1043
 : 1041
 : 1043
 : 1046
 : 1041
 : 1045
 : 1047
 : 1041

[illegible][illegible]

7.3 ARTICLE 1: A canadian isolate of hepatitis *delta* virus.

Véronique Poisson, Daniel B. Ménard, Éric Frost et Jean-Pierre Perreault

Département de Biochimie, de Médecine, de Microbiologie et d'Infectiologie
Université de Sherbrooke
Sherbrooke, Qué., Canada, J1H 5N4

Mai 1999

Article accepté pour publication dans: *Canadian Journal of Gastroenterology*

A Canadian isolate of hepatitis *delta* virus.

running title: HDV RNA sequence

Véronique Poisson BSc¹, Daniel B. Ménard MD², Éric Frost PhD³, Jean-Pierre Perreault PhD¹

Département de Biochimie¹, Département de Médecine², Département de Microbiologie et Infectiologie³, Faculté de médecine, Université de Sherbrooke, Sherbrooke, Québec, J1H 5N4, Canada.

°

May 14th 1999

To whom correspondence should be addressed: Dr Jean-Pierre Perreault, Département de biochimie, Université de Sherbrooke, Sherbrooke, Québec, J1H 5N4. Telephone: (819) 564-5310, fax: (819) 564-5340, e-mail: jperre01@courrier.usherb.ca

ABSTRACT

Hepatitis *delta* virus (HDV) is an infectious agent that propagates in hepatocytes only in the presence of hepatitis B virus, causing fulminant or chronic hepatitis with liver cirrhosis. HDV is a 36 nm particle which includes a circular RNA genome of 1,7 kilobases with extensive internal complementary that allows it to fold into a rod-like structure. The relationships among genotypes, sequence variability, geographical distribution and disease severity of HDV remain unknown. Consequently, we determined the complete nucleotide sequence of an HDV isolated from a Canadian individual. The viral RNA from serum was amplified using reverse transcription coupled to polymerase chain reaction amplification. The resulting cDNA was cloned and sequenced. Sequence analysis revealed that this new isolate contained 1672 nucleotides corresponding to genotype I, which has a worldwide distribution. Sequencing of four independent clones revealed 17 substitutions, corresponding to an overall sequence variability of 1%. Surprisingly, 7 mutations were found in the 48 nucleotide region located between the two highly conserved self-catalytic motifs. This is the first demonstration that many substitutions are identified in this region of HDV, and prompts us to define it as a hypervariable region.

Key Words: Viral hepatitis, Hepatitis *delta* virus, genotype, sequence variability, RT-PCR, sequence analysis

RÉSUMÉ

Le virus de l'hépatite *delta* (VHD) est un agent infectieux qui se propage dans les hépatocytes seulement en présence du virus de l'hépatite B, causant ainsi des hépatites fulminantes et chroniques avec des cirrhoses du foie. VHD est une particule de 36 nm contenant un génome d'ARN circulaire de 1,7 kilobases qui présente beaucoup de complémentarité interne le conduisant à se replier en une structure semblable à une tige. Les relations entre génotypes, degré de variabilité des séquences, distribution géographique, et sévérité de la maladie sont des aspects encore incompris. Par conséquent, nous avons déterminé la séquence nucléotidique complète d'un VHD isolé d'un individu canadien. L'ARN viral d'un sérum fut amplifié en utilisant une transcription inverse couplée à une réaction de polymérisation en chaîne. L'ADNc résultant fut ensuite cloné et séquencé. L'analyse de la séquence a révélé que ce nouvel isolat contenant 1672 nucléotides correspond au génotype I, lequel a une distribution mondiale. Le séquençage de quatre clones-indépendants a révélé 17 substitutions, correspondant à une variabilité de séquence moyenne de 1%. Il est intéressant de noter que 7 de ces mutations ont été retrouvées dans la petite région constituée de 48 nucléotides localisée entre les deux motifs autocatalytiques hautement conservés. C'est la première fois qu'autant de substitutions sont identifiées dans cette région de VHD, nous suggérant de la définir comme une région hypervariable.

Mots Clés: Hépatite virale, Virus de l'hépatite *delta* , génotype, variabilité de séquence, RT-PCR, analyse de séquence

INTRODUCTION

Hepatitis *delta* virus (HDV) is an infectious agent that infects and propagates in hepatocytes exclusively in the presence of hepatitis B virus (HBV), causing fulminant or chronic hepatitis with liver cirrhosis (reviewed in 1,2). The symptoms associated with both acute and chronic HDV infection are frequently more severe than those seen with HBV alone, and appear to differ in severity depending upon the geographic location. HDV is a 36 nm particle and is unique among animal viruses in that it contains a single-stranded circular RNA genome of 1,7 kilobases with extensive internal complementarity that allows it to assume a rod-like structure (1,2). These features are reminiscent of those seen in the plant RNA pathogens known as viroids. Moreover, both HDV and viroids have RNA genomes that undergo RNA-dependent RNA replication via a rolling circle mechanism that is DNA-independent (i.e. involves only RNA intermediates) (2). The HDV genome is composed of two domains known as the left- and right-hand domains (Fig. 1). The left-hand domain comprises the viroid-like region and includes self-cleaving sequences for RNAs of both genomic and antigenomic polarities. These two self-cleaving RNA motifs are crucial for HDV replication. The right-hand domain includes two regions: the upper strand encodes the mRNA for a single protein, the *delta* antigen (HDAg); while the lower strand is a variable domain to which no function has been attributed. HDAg, which is encoded by a mRNA of sequence corresponding to the antigenomic polarity, exists as two distinct species: a large HDAg (L-HDAg, 27 kDa, 214 amino acids) and a small HDAg (S-HDAg, 24 kDa, 195 amino acids). The L-HDAg is produced following an editing event that alters the first stop codon to a tryptophan codon thereby extending the translation to the next stop codon. S-HDAg acts as a transcriptional factor during HDV RNA replication, whereas L-HDAg inhibits replication and is required for virus assembly.

Comparison of HDV sequences indicated the existence of three distinct genotypes (3,4). It has been observed that genotypes II and III are 80% and 65% homologous, respectively, with genotype I. Genotype I is widespread in Italy, the United States, Taiwan, Nauru, France, Lebanon and China. The disease pattern associated with infection by genotype I HDV, appears to be highly variable, ranging from severe to mild. Genotype II has been isolated only in Japan and Taiwan, and is associated with a milder disease than that caused by genotype I infection. Genotype III has been isolated only in northern parts of South America where the disease is the most severe. Moreover, a sequence microheterogeneity was observed in patients. The role of HDV genetic variations in the severity of hepatitis *delta* is yet clear, but some studies suggest that they are linked (3,5). Conversely, the relationship between HDV genotypes and the degree of HDV sequence variability in relation to geographical distribution and disease severity remains an unclear issue. To address this question we determined the nucleotide sequence of a Canadian HDV isolate, and investigated the sequence variability throughout the whole genome (as compared to most studies that are limited to the protein-coding region).

PATIENT AND METHODS

This study was approved by the ethical review board of the Université de Sherbrooke, Centre de Recherche Clinique. The serum sample from which the HDV RNA was isolated was collected from a 50 year old Canadian male resident of the province of Quebec. This patient suffered from a severe chronic hepatitis which was rapidly evolving into cirrhosis and liver failure. The serum was previously analyzed for hepatitis B surface antigen, IgM antibody to hepatitis B core antigen (IgM anti-HBc), total antigen to hepatitis *delta* (anti-HD), IgM antibody to hepatitis *delta* (IgM anti-HD), and antibody to

hepatitis C virus (HCV) (Elisa II). The diagnosis of hepatitis C was also confirmed by polymerase chain reaction (PCR).

The serum (50 μ L) was subjected to proteinase K digestion, and then phenol / chloroforme (1 vol:1 vol) extracted. The mixture of cellular and viral RNA was precipitated with isopropanol, washed with 70% ethanol, and dissolved in 50 μ L DEPC treated water. A strategy of reverse transcription (RT) coupled to PCR was used to amplify four fragments of the HDV genome. The sequences of the primers used for RT-PCR are shown in Table 1. Only one oligonucleotide (AS1100) was required for the RT reaction. Briefly, one tenth of the extracted RNA and 100 pmol of primer AS1100 were annealed in a final volume of 18 μ L by incubating at 70°C for 5 min. The reaction mixture was then adjusted to 30 μ L such that it contained 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM dithiotreitol, 0.67 mM of each dNTP, 30 units of the ribonuclease inhibitor RNA Guard (Pharmacia Biotech) and 200 units of SuperScript™ II reverse transcriptase (Gibco BRL), and incubated at 42°C for 50 min. The pairs of oligonucleotides shown in Figure 2A were used in the subsequent PCR steps. One fifth of the cDNA pool (6 μ L) was added to a PCR mixture containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM MgCl₂, 0.2 mM of each dNTP, 100 pmol of both sense and antisense primers and 1 unit *Taq* DNA polymerase in a final volume of 100 μ L. Thirty amplification cycles (one min at 94°C, 1.5 min at a temperature between 52°C and 65°C depending of the primer set used, and 1 min at 72°C) were performed. After amplification, the mixtures were extracted with phenol / chloroforme, and 20 μ L aliquots analyzed by 1.5 % agarose gel electrophoresis (see Fig. 2B). Gel slices containing the four PCR products (838 bp, 610 bp, 606 bp and 125 bp in length) were isolated, and the DNA extracted, precipitated and washed with ethanol. The resulting DNA fragments were cloned by taking advantage of the property of *Taq* DNA polymerase of adding an

adenosine at the 3' ends of the PCR products. The PCR amplified fragments were ligated in a "sticky end" fashion to linearised pCR^R 2.1 vector, which possesses an extra thymidine residue at each 5' end, as recommended by the manufacturer (TA cloning kit; Invitrogen) (6). Four white colonies for each PCR product were chosen after plating on petris containing ampicillin, IPTG and X-Gal, and were sequenced by the dideoxyribonucleotide chain termination method using the T7 sequencing kit (Pharmacia Biotech). Sequencing reactions were carried out using the M13 universal primer and custom oligonucleotides derived from the HDV genome as shown in Table 1.

RESULTS AND DISCUSSION

In order to sequence the entire genome of an HDV isolate from a Canadian patient, we used a strategy of reverse transcriptase (RT) for the production of a complementary DNA molecule coupled to amplification by polymerase chain reaction (PCR). Although HDV is a highly structured RNA molecule, we found that the oligonucleotide AS1100 was sufficient to produce a full length cDNA molecule. Oligonucleotides complementary to various HDV regions were tested; however, they failed to produce full length cDNAs. Oligonucleotide AS1100 annealed to a region of the HDV genome that is highly conserved in all genotypes, and may therefore be considered as an original and universal primer for the production of full length HDV cDNA. PCR amplification of the cDNA was performed using four distinct pairs of primers that produced overlapping fragments spanning the entire genome (see Table 1 and Fig. 2A). In most previous studies five or more sets of primers were required to amplify the complete genome (7-10). The primers used in this work were selected from highly conserved regions of the HDV genome as determined by a sequence alignment of all known isolates (11), and resulted in optimal RT-PCR reactions. Primers located in highly conserved regions should be more reliable,

and therefore more sensitive than others in detecting HDV sequences. Furthermore, the approach of using several regions, as used here, facilitated the subsequent sequencing step, as well as reducing the probability of PCR derived sequence artifacts. The four primer pairs used here produced the predicted fragments (125-838 basepairs) as revealed by agarose gel electrophoresis (Fig. 2B).

Each RT-PCR product was extracted from the agarose gel, cloned, and sequenced. The complete sequence of this HDV isolate is shown in figure 2C (GenBank accession number: AF098261). The sequence data reveals an RNA molecule 1672 bases in length with characteristics similar to other HDV species. For example, one open reading frame encodes a functional protein with all of the features required for its activity, and possesses the appropriate S-HDAg stop codon of UAG which can be edited so as to produce the L-HDAg form. The alignment of this sequence with the other 22 complete sequences reported to date indicates that this HDV isolate is of genotype I. The sequences of the four chosen clones showed the existence of some variability in the patient (Fig. 2C; i.e. for any given fragment, none of the clones were completely identical). Sequence divergence was detected over the entire genome, reflecting the selective pressures. For example, the protein-coding region (HDAg) was highly conserved, while the variable region had several mutations. In total, mutations were detected at 17 positions, corresponding to 1% variation. Among the observed mutations, five appeared for the first time (underlined in Fig. 2C). This microheterogeneity is probably real and not the result of PCR artifacts because this divergence rate is much higher than the reported error frequency of *Taq* polymerase (12). Rather, it likely reflects the actual microheterogeneity of HDV RNA within an individual. This phenomenon may contribute to the diversity seen in the clinical picture of HDV infection. More interestingly, 7 out of the 17 mutations were detected in the small region between both the genomic and antigenomic self-cleaving domains. Although no function was attributed to this region, this is the first

time that numerous mutations have been observed in this small region of 48 bases. This may have significant consequences on the molecular and cellular biology of HDV, although this question remains to be addressed. These results suggest that this region maybe considered as being hypervariable.

CONCLUSION

Sequence analysis of an HDV species isolated from the serum of a Canadian patient indicated that it is of genotype I. The sequence showed a divergence, evenly distributed over the entire genome, from the previously known HDV isolates. These results confirm that HDV species isolated from different geographical areas are heterogeneous. The sequence heterogeneity of HDV RNA might, to some extent, explain the variations observed in the degree of the severity of *delta* hepatitis and in the transmissibility of the virus. However, the role of HDV genetic variations in determining the severity of hepatitis *delta* remains unclear. Other factors contributing to the clinical manifestations of HDV infections could be the nature and condition of the associated HBV infections. Further analysis of the worldwide distribution of different HDV sequences, and of sequence variations within HDV isolates, will be important for the further evaluation of the pathogenic and epidemiologic aspects of this infectious agent.

ACKNOWLEDGEMENTS

We thank Michael Dionne for its technical assistance and Dr Sirinart Ananvoranich for critical reading of the manuscript. This work was supported by a grant from Medical Research Council (MRC) of Canada to J.P.P.. V.P. is recipient of a studentship from Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR). J.P.P. is an MRC scholar.

REFERENCES

1. Lai MMC. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* 1995;64:259-86.
2. Poisson F, Roingeard P, Goudeau A. Le virus de l'hépatite delta: un mode de réplication bien singulier. *Médecine/Sciences* 1995;10:1379-87.
3. Casey JL, Brown TL, Colan EJ, Wignall FS, Gerin JL. A genotype of hepatitis D virus that occurs in Northern South-America. *Proc. Nat. Acad. Sci. USA* 1993;90:9016-20.
4. Wu JC, Chiang TZ, Sheen IJ. Characterization and phylogenetic analysis of a novel hepatitis D strain discovered by restriction fragment length polymorphism analysis. *J. Gen. Virol.* 1998;79:1105-13.
5. Wu JC, Choo KB, Chen CM, Chen TZ, Huo TI, Lee SD. Genotyping of hepatitis D virus by restriction-fragment length polymorphism and relation to outcome of hepatitis D. *Lancet* 1995;346:939-41.
6. Lafontaine DA, Mercure S, Perreault JP. Identification of a Crohn's disease specific transcript with potential as a diagnostic marker. *Gut* 1998;42:878-882.
7. Chao YC, Chang MF, Gust I, Lai MMC. Sequence conservation and divergence of hepatitis virus RNA. *Virology* 1990;178:384-92.
8. Chao YC, Lee CM, Tang HS, Govindarajan S, Lai MMC. Molecular cloning and characterization of an isolate of hepatitis delta virus from Taiwan. *Hepatol.* 1991;13:345-352.
9. Lee CM, Bih FY, Chao YC, Govindarajan S, Lai MMC. Evolution of hepatitis delta virus RNA during chronic infection. *Virology* 1992;188:265-273.
10. Lee CM, Changchien CS, Chung JC, Liaw YF. Characterization of a new genotype II hepatitis delta virus from Taiwan. *J. Med. Virol.* 1996;49:145-154.

11. Lafontaine AD, Mercure S, Perreault J.-P. Update of the viroid and viroid-like RNA sequence database: addition of the hepatitis *delta* RNA section. *Nucleic Acids Res.* 1997;25:123-25.
12. Shakil AO, Hadziyannis S, Hoofnagle JH, Di Bisceglie AM, Gerin JL, Casey JL. Geographic distribution and genetic variability of hepatitis *delta* virus genotype 1. *Virology* 1997;234:160-67.
13. Makino S, Chang MF, Shih CK, et al. Molecular cloning and sequencing of human hepatitis virus RNA. *Nature* 1987;329:343-46.
14. Kuo MY, Chao M, Taylor J. Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of delta antigen. *J. Virol.* 1989;63:1945-50.

FIGURE LEGEND

FIGURE 1) Schematic diagram of the hepatitis *delta* virus genome. The RNA genome of HDV folds into a rod-like structure. The left-hand domain (boxed) is reminiscent of the plant viroids, and includes self-cleaving sequences of both genomic and antigenomic polarities (striped rectangles). The self-cleavage sites are indicated by arrows. The open reading frame for both forms of *delta* antigen (black box) is on the upper strand of the right-hand domain, while the lower strand is the variable domain.

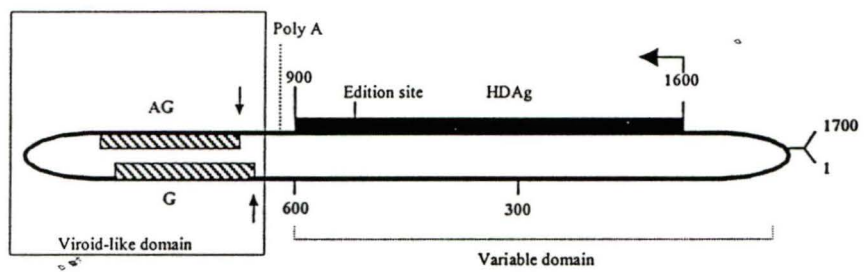
FIGURE 2) Molecular cloning and sequencing of an HDV isolate. **A** Schematic representation of the fragments amplified by PCR. The arrow indicates the orientation of the genomic RNA. **B** PCR products were purified on 1.5% agarose gels. Lanes 1-4 are negative controls performed using the set of primers S700-AS1200: lane 1, RT-PCR performed in the absence of RNA as template; lane 2, RT reaction performed in the absence of the antisense primer, which was only added for the subsequent PCR reaction; lane 3, RT-PCR using the sense primer to perform the RT reaction; and, lane 4, omission of the RT reaction. Similar controls were performed with all four sets of primers (data not shown). Lanes 5, 7, 9 and 11 are positive controls performed using a cloned HDV dimer sequence (pSVL (D3), generously provided by Dr John Taylor; 14), while lanes 6, 8, 10 and 12 are the RT-PCR amplifications performed using the RNA extracted from the Canadian patient as template. The primer sets used for the PCR amplifications are: lanes 5 and 6, S700 / AS1200; lanes 7 and 8, S1100 / AS275; lanes 9 and 10, S275 / AS800; and, lanes 11 and 12, S200 /

AS300. Lanes M₁ and M₂ are the DNA size markers, HaeIII digested ϕ X174 (Promega) and 1Kb DNA ladder (Gibco BRL), respectively. **C** RNA sequence of the HDV isolate. The genomic site corresponding to the start codon and the orientation of HDAg are indicated. The vertical lines indicate the termination sites for both the S- and L-HDAg. The arrows at positions 681 and 895 indicate the self-cleaving sites of the genomic and antigenomic strands, respectively. The nucleotide variations observed are indicated in parenthesis. The underlined nucleotides correspond to positions that are shown to vary for the first time. The sequence in the box corresponds to the small hypervariable region, located between the self-cleaving sequences, that we identified.

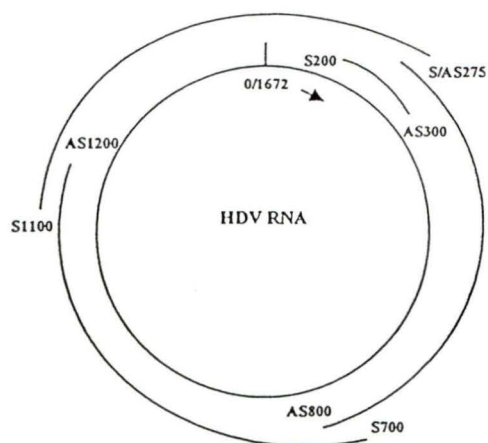
TABLE 1
Synthetic oligonucleotide primers.*

Name	Sequence	Position †
Sens primers		
S200	5'dCGATCCGAGGGGCCAAC3'	202-219
S275	5'dCACTCCGGCCCGAAGGGTTG3'	236-255
S700	5'dTGGCCGGCATGGTCCCAGCC3'	685-704
S1100	5'dGCCGCCCGGGCGCTCC3'	1101-1118
Antisens primers:		
AS275	5'dCAACCTTCGGGCCGGAGTG3'	255-236
AS300	5'dGCTGAAGGGTCTCTGG3'	324-304
AS800	5'dCGCACGTCCACTCGGATG3'	844-827
AS1100	5'dGGAAAAGAAGAGTAGCCG3'	1180-1163
AS1200	5'dGAAGGAAGGCCCTCGAGAACAAG3'	1288-1266
Sequencing primers		
F606	5'dGAAGCGGATCGAGGGGAGC3'	1133-1114
F606N	5'dCTTCTCTTCGGGTCGG3'	911-894
F836N	5'dGCTTCTGGGAGTAGTTTC3'	106-88
F836Z	5'dGAGGACGAAAATCCCTGG3'	1466-1449

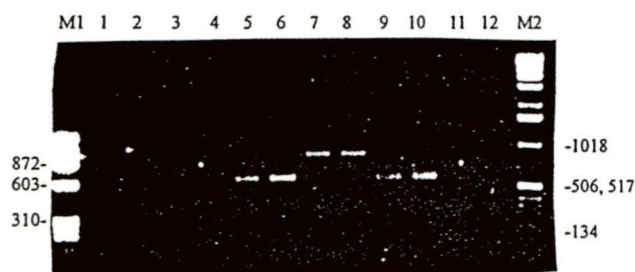
*Primers were designed using a HDV sequences alignment as described in 'Result and Discussion';†The position of primers is according to Makino et al. (13).



A



B



C

¹CCTGAGCCAAGTTCCGAGCGAGGAGACGCGGGGGGAGGATCAGCTCCCAGAGGGGATGT
⁶¹CACGGTAAAGAGCATTGGAACGTCGGAGAACTACTCCCAAGAAGCAAAGAGAGGTCTTA
¹²¹GGAAGCGGACGAGATCCCCACAACGCCGAGAATCTCTGGAAGGGGAAAGAGGAAGGTGG
¹⁴¹AAGAAAAGGGGCGGGCTCCCGATCCGAGGGGCCAACTTAATCTCCAGATCTGGAGAG
²⁴¹CACTCCGGCCCGAAGGGTTGAGAGTACCCAGAGGGAGGAAGCCACACGGAGTAGAACAG
³⁰¹AAAAATCACCTCCAGAGACCCCTTCAGCGAACAGAGAGCGCATCGCAGAGGGAGTAGAC
³⁶¹CATAGCGATAGGAGGGGATGCTAGGAGTTAGGGGAGACCAAGCGAGGAGAAAGTAAAGAA
⁴²¹GCAACGGGGCTAGCAGGTGGGTGTTCCGCCTCCGAGAGACGAGTGAGGCTTATCCCGGG
⁴⁸¹GAACTCGACTTATCGTCCCCACATAGCAGGACCCCGGACCCCTTCGAAAGGTACCGGA
⁵⁴¹GGGGGTGACTCGAACACCGGGGACAGTGAGGCCATGGGATGCTCTTCCCGATGCTCGA
⁶⁰¹TTCCGACTCCCCCAAGGGTCGCCAGGAATGGCGGGACCCACTCTGCAGGGTCCGCG
⁶⁶¹TTCCATCCTTTCTTACCTGATGGCCGGCATGGTCCAGCCTCCTCGTGCGCGGGCTGGG
⁷²¹CAACATTCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCCACAAACCTCTCTAG
⁷⁸¹CTTCCAGAGAGAAGCGAGAGGAAAGTGGCTCTCCCTTAGCCATCCGAGTGGAGCTGCGT
⁸⁴¹CCTTCTTCGGATGCCAGGTCCGACCGCGAGGAGGTGGAGATGCCATGCCGACCCGAAGA
⁹⁰¹GGAAAAGGACGCGAGACGCAAACTGTGAGTGGAACCCGCTTTATCTACTGGGGTCG
⁹⁶¹ACAACTCTGGGAGAGAAGGGAGGGTCCGCTGGGAAGAGTATATCTATGGGAATCCCTG
¹⁰²¹GCTTCCCTGATGTCCAGTCCCTCCCGGTCGAGAGAAGGGGACTCCGGGACCCCTTG
¹⁰⁸¹ATGCTGGGACGAAGCCGCCCGGGCGCTCCCTCGATCCGCTTCGAGGGGTTTACA
¹¹⁴¹CCCCCAACTGCGGGCGGGCTACTCTTCTTTCCCTTCTCTGCTTCTCCTCGGTCAACCTT
¹²⁰¹TAAGTTCTTCTTCTCTCTTCTGCTGAGGCTCTTCCCCCGCGGATAGCTGCTTCTCT
¹²⁶¹TGTTCTCGAGGGCTTCTTCTGTCGGTGATCCTGCCTCTCTTGTGCGTGAATCCTCCCC
¹³²¹TGAGAGGCCTCTTCCAGGTCCGAGTCTACCTCCATCTGGTCCGTTCCGGGCCCTCTTCG
¹³⁸¹CCGGGGAGCCCCCTCTCCATCCTTATCCTTCTTCCGAGAATCCTTTGATGTTCCCCA
¹⁴⁴¹GCCAGGGATTTTCGTCCTCTATCTTCTTGAGTTTCTTCTTGTCTCCGGAGGTCTCTCT
¹⁵⁰¹CGAGTTCCTCTAACTTCTTCTTCCGGCCACCCACTGCTCGAGGATCTCTCTCTCCCTC
¹⁵⁶¹CGCGGTTCTTCTCGACTCGGACCGGCTCATCTCGGCTAGAGGGCGAGTCTCTAGTACT
¹⁶²¹CTTACTCTTTCTGTAAAGAGGAGACTGCTGGACTCGCCGCCGAGCCCGAG

7.4 ARTICLE 2: The viroid and viroid-like RNA database.

Daniel A. Lafontaine, Partick Peschênes, Frédéric Bussière,
Véronique Poisson et Jean-Pierre Perreault

Département de Biochimie
Université de Sherbrooke
Sherbrooke, Qué., Canada, J1H 5N4

Septembre 1998

Article publié dans: *Nucleic Acids Research*, 1999, Vol. 27, No. 1

The viroid and viroid-like RNA database

Daniel A. Lafontaine, Patrick Deschênes, Frédéric Bussière, Véronique Poisson and Jean-Pierre Perreault*

Département de Biochimie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada

Received August 28, 1998; Accepted September 1, 1998

ABSTRACT

This is an online database to facilitate research on viroid, viroid-like RNAs and human hepatitis *delta* virus (vHDV) by presenting a large number of sequences and related data in a comprehensive and user-friendly format (e.g. position of their self-catalytic domains, open reading frame of the vHDV, prediction of the most stable secondary structures, etc.). Most of these RNA species share a common proposed replication pattern known as a DNA-independent rolling circle mechanism. Together, these species form the 'brotherhood' of the smallest known auto-replicable RNAs. This online database is available on the World Wide Web at <http://www.callisto.si.usherb.ca/~jpperra>

INTRODUCTION

The initial, 1996, version of this database focused on sequences from viroids, which are small, circular, single-stranded RNA infecting plants, and plant satellite viroid-like RNAs (1). In the 1997 update, we included a section on the human hepatitis *delta* virus (vHDV) (2). The 1998 version was reorganized to be more user-friendly and easier to access the database (3). The four sections (viroids, satellite RNAs, HDV and others) of the databank are shown in a frame which is always available and allows easy access to any subdivision (Table 1). In addition to the update of the existing database, the 1999 version offers several novel features as described below.

DESCRIPTION

The choice of a section (Table 1) will lead to the second level which is a summary table for each sequence subdivision. The viroids, satellite RNAs and other RNAs are divided according to the nature of the compiled RNA species. The choice of either a viroid or viroid-like RNA species in a summary level will lead to the third level. Briefly, each species of RNA appearing in the database is listed by its complete name and number of sequence variants. This is followed, for each species, by a complete listing of the sequence variants and their assigned nomenclature. The identification of species variants is based on its usual acronym followed by a number. The procedure for sequence identification and information compiled was presented previously (1). Additional data of each entry include their accession numbers in

Table 1. Summary of the database sections

VIROIDS
· ASBV-type (group A)
· PSTV-type (group B)
· PSTVd subgroup
· CCCVd subgroup
· HSDV subgroup
· ASSVd subgroup
· CbVd subgroup
SATELLITE RNAs
· Lutecovirus
· Nepovirus
· Sobemovirus
· Other satellites
View sequence alignments
vHDV RNA SEQUENCES
· Complete genome sequences
· Partial related RNA sequences
View sequence alignments
OTHERS

sequence library file servers, bank loci (when available), number of nucleotides (total and by type), complete publication information and the sequence in blocks of 10 nucleotides. For several sequences, a table shows structural features, for example the position of the conserved sequences of the self-catalytic domain (hammerhead, hairpin, delta and VS RNA). In addition, secondary structure prediction of the most likely ancestral variant of each entry (except for vHDV and VS RNA) was derived using the RNAfold structure prediction package. These predicted structures are appended to the database in connect file format to allow easy manipulation (fourth level).

Unlike the viroid section, the vHDV section comprises all available sequences, irrespective of their completeness, since the majority of the partial sequences are informative (i.e. they correspond to the sequence of either the open reading frame or the ribozymes). Both HDV complete and partial sequences were arbitrarily subdivided into several parts in order to accelerate the

* To whom correspondence should be addressed. Tel: +1 819 564 5310, Fax: +1 819 564 5340, Email: jpperra01@commerce.usherb.ca

display. For the partial vHDV sequence, a 'p' preceding the specific number is attributed to partial sequence (2). From the 1998 version of the database, sequence alignments of all complete vHDV nucleotide sequences and all amino acid sequences of the vHDV mRNA antigen have been available through the summary tables (second level). This database provides an excellent reference point for further phylogenetic and structure-function studies of these RNA species.

1999 VERSION

Today, more than 450 sequences are now part of this database comprising data from 25 viroid species, nine species of plant satellite viroid-like RNAs, four related species of RNA and vHDV (23 complete and 165 partial). The database is now used daily by several researchers, and all user comments are truly appreciated by the authors. Clearly, the development of the database depends on input from members of the research community. The modifications and additions performed in this new version are in response to constructive comments; for example: (i) in agreement with the recent review of Flores and collaborators (4), the group B viroids were clustered in five subgroups according to the sequence composing their central conserved regions (CCR) (Table 1); (ii) within the vHDV RNA sequence sections, alignment of all sequences for the self-catalytic domains as well as for the transcriptional promoter are in progress and will be added soon; (iii) within the satellite RNAs section, the alignment of hammerhead self-catalytic domains is in progress and will be added soon. Subsequently, a phylogenetic reconstruction of the viroids and plant satellite RNA including hammerhead sequence will soon be available; etc.

COMPLETENESS, ACCURACY AND AVAILABILITY OF THE DATA

This databank is available on the World Wide Web browser (e.g. Netscape Navigator) at the URL <http://www.callisto.si.usherb.ca/~jpperra>. Floppy disks (readable on microcomputers operating under MS-DOS or Macintosh) and hard copy version are also available upon request only for those without electronic access to the database. This database is updated as soon as sequences become available. Users of the viroid and viroid-like RNA database should cite this publication and are encouraged to provide corrections, new information or other information for inclusion in the database via electronic mail (jperre01@courier.usherb.ca). The authors would appreciate being informed of any omitted sequences or errors in the data set. We will correct any errors.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Natural Sciences and Engineering Research Council (NSERC), the Medical Research Council (MRC) of Canada, and in part by a scientific team grant from Fonds pour la Formation des Chercheurs et l'Avancement de la Recherche du Québec (FCAR) to J.-P.P. F.B. is a recipient of predoctoral fellowships from NSERC while D.L., P.D. and V.P. are recipients of the Fonds de la Recherche en Santé du Québec (FRSQ). J.-P.P. is an MRC scholar.

REFERENCES

- 1 Bussi re, F., Lafontaine, D. and Perreault, J.-P. (1996) *Nucleic Acids Res.*, **24**, 1793-1798.
- 2 Lafontaine, D., Mercure, S. and Perreault, J.-P. (1997) *Nucleic Acids Res.*, **25**, 123-125.
- 3 Lafontaine, D.A., Mercure, S., Poisson, V. and Perreault, J.-P. (1998) *Nucleic Acids Res.*, **26**, 190-191.
- 4 Flores, R., Di Serio, F. and Hernandez, C. (1997) *Seminars Virol.*, **8**, 65-73.

7.5 **ARTICLE 2:** The viroid and viroid-like RNA database.

Daniel A. Lafontaine, Stéphane Mercure, Véronique Poisson
et Jean-Pierre Perreault

Département de Biochimie
Université de Sherbrooke
Sherbrooke, Qué., Canada, J1H 5N4

Août 1997

The viroid and viroid-like RNA database

Daniel A. Lafontaine, Stéphane Mercure, Véronique Poisson and Jean-Pierre Perreault*

Département de Biochimie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada

Received August 6, 1997; Accepted August 7, 1997

ABSTRACT

The viroid and viroid-like RNA database is a compilation of all natural sequences published in journals or available from the GenBank and EMBL nucleotide sequence libraries. Several information regarding these RNA species such as the position of their self-catalytic domains and the open reading frame of the human hepatitis *delta* virus are provided. The database also includes a determination of the likely ancestral sequence of most species and a prediction of the most stable secondary structures of these sequences. This online database is available on the World Wide Web (<http://www.callisto.si.usherb.ca/~jpperra>). It should provide an excellent reference point for further phylogenetic and structure-function studies of these RNA species.

DESCRIPTION

We have developed an online database in order to facilitate viroid, viroid-like RNAs and human hepatitis *delta* virus (vHDV) research by presenting a large number of sequence and related data in a comprehensive and user-friendly format. Most of these RNA species share a common proposed replication pattern known as a DNA-independent rolling circle mechanism. Together, these species form the 'brotherhood' of the smallest known auto-replicable RNAs. The initial version of this database focused on sequences from viroids, which are small, circular, single-stranded RNA infecting plants, and plant satellite viroid-like RNAs (1). In a subsequent update, we included a section focusing on the human hepatitis *delta* virus (vHDV) (2). Unlike the viroid section, the vHDV section comprises all available sequences, irrespective of their completeness, since the majority of the partial sequence are informative (i.e. they correspond to the complete sequence of either the open reading frame or the ribozymes).

In this manuscript, we describe an update of this database which to date includes more than 400 sequences from 36 different species. In addition to the update of the existing database, this new version offers a user-friendly structure which facilitates and accelerates the access to the database. The four sections of the databank are now shown in a frame which is always available and allows easy access to any subdivision:

•viroids •satellite RNAs •HDV •others

The choice of a section will lead to the second level which is a summary table for each sequence subdivision (Fig. 1). The viroids, satellites RNAs and other RNAs are divided according to the nature

of the compiled RNA species. In contrast, the HDV section is subdivided on the basis of sequence completeness (i.e. complete and partial). Both of HDV complete and incomplete sequences were arbitrarily subdivided in two parts in order to accelerate the display.

The choice of either a viroid or viroid-like RNA species in a summary level will lead to the third level. Briefly, each species of RNA appearing in the database is listed by its complete name and number of sequence variants. This is followed, for each species, by a complete listing of the sequence variants and their assigned nomenclature. The identification of species variants is based on its usual acronym followed by a number. The procedure for sequence identification and information compiled was presented previously (1). For the partial vHDV sequence, a 'p' preceding the specific number is attributed to partial sequence (2). Additional data in each entry includes their accession numbers in sequence library file servers, bank loci (when available), number of nucleotides (total and by type), complete publication information, and the sequence in 10 nucleotide blocks. For several sequences, a table lists structural features as the position of the conserved sequences of the self-catalytic domains, completes the entry. In addition, a secondary structure prediction of the most likely ancestral variant of each entry (except for vHDV and VS RNA) was derived using the RNAfold structure prediction package. These predicted structures are appended to the database in connect file format (fourth level). The analysis of the viroid and viroid-like RNA compilation (e.g. classification, secondary structure prediction, phylogenetic identification of the likely ancestral variant, etc.) has been published previously (1).

Two sequence alignment files will be also available, in early 1998, through the summary tables (second level). These sequence alignment will be performed using: (i) all complete vHDV nucleotide sequences; and (ii) all amino acid sequences of the vHDV mRNA antigen. These alignments as well as the database itself should provide an excellent reference point for further phylogenetic and structure-function studies of these RNA species.

AVAILABILITY

This databank is available on the World Wide Web browser (e.g. Netscape Navigator) at the URL <http://www.callisto.si.usherb.ca/~jpperra>. Floppy disks (readable on microcomputers operating under MS-DOS) and in hard copies version are also available upon request for those without electronic access to the database. This database is updated bimonthly. Users of the viroid and viroid-like RNA database should cite this publication and are encouraged to provide corrections, new information or other information for inclusion in the database via electronic mail (jp.perre@courrier.usherb.ca).

* To whom correspondence should be addressed. Tel. +1 519 561 5310 Ext. 41519 561 5330. E-mail: jpperra@courrier.usherb.ca

Compilation of viroids

RNA species	Abbreviation	Number of sequences	RNA species	Abbreviation	Number of sequences
Viroids					
ASBVd-type (group A)					
<u>Avocado sunblotch viroid</u>	ASBVd	19	PSTVd-type (group B)		
<u>Peach latent mosaic viroid</u>	PLMVd	6	ASSVd group (subgroup B2)		
PSTVd-type (group B)			<u>Australian grapevine viroid</u>	AGVd	1
PSTVd group (subgroup B1)			<u>Apple dimple fruit viroid</u>	ADFVd	1
<u>Coconut cadang-cadang viroid</u>	CCCVd	7	<u>Apple scar skin viroid</u>	ASSVd	4
<u>Citrus exocortis viroid</u>	CEVd	33	<u>Citrus bent leaf viroid</u>	CBLVd	3
<u>Columnnea latent viroid</u>	CLVd	3	<u>Citrus viroid species III</u>	Cvd-III	2
<u>Chrysanthemum stunt viroid</u>	CSVd	4	<u>Coleus blumei viroid I</u>	CoVd	9
<u>Coconut tinangaja viroid</u>	CTVd	2	<u>Grapevine yellow speckle viroid</u>	GYSVd	34
<u>Citrus viroid species</u>	CVdIV	1	<u>Grapevine IIB viroid</u>	GIBVd	1
<u>Hop latent viroid</u>	HLVd	1	<u>Pear blister canker viroid</u>	PBCVd	1
<u>Hop stunt viroid</u>	HSVd	32			
<u>Iresine viroid</u>	IRVd	1			
<u>Mexican papaya viroid</u>	MPVd	9			
<u>Potato spindle tuber viroid</u>	PSTVd	29			
<u>Tomato apical stunt viroid</u>	TASVd	3			
<u>Tomato planta macho viroid</u>	TPMVd	1			

Compilation of satellite RNAs

RNA species	Abbreviation	Number of sequences
Satellite RNA		
Luteovirus		
<u>Barley yellow dwarf virus satellite RNA</u>	vBYDV	1
Nepovirus		
<u>Arabidopsis mosaic virus satellite RNA</u>	sARMV	1
<u>Chicory yellow mottle virus satellite RNA S1</u>	sCYMV-S1	1
<u>Lucerne transient streak virus satellite RNA</u>	vLTSV	3
<u>Tobacco ringspot virus satellite RNA</u>	sTobRV	2
Sobemovirus		
<u>Subterranean clover mottle virus satellite RNA</u>	vSCMoV	2
<u>Solanum nodiflorum mottle virus satellite RNA</u>	vSNMV	1
<u>Velvet tobacco mottle virus satellite RNA</u>	vTMoV	2

Compilation of HDV sequences

RNA species	Number of sequences
Complete sequences..... <u>part I, part II</u>	20
Partial sequences..... <u>part I, part II</u>	50

Compilation of other related RNAs

RNA species	Abbreviation	Number of sequences
Other related RNA		
<u>Carnation stunt associated viroid</u>	rCarSV	2
<u>Newt satellite 2 transcript</u>	rNS2T	1
<u>Neurospora VS RNA</u>	VS RNA	1

Figure 1. Summary tables of the four sections.

ACKNOWLEDGEMENTS

Authors would like to thank Annie Dumontier and Jacques Lehoucq for technical assistance in the development of this update. This work was supported in part by grants from the Natural Sciences and Engineering Research Council (NSERC) and by the Medical Research Council (MRC) of Canada to J.-P.P., E.B. and D.L., are recipients of post-doctoral fellowships from NSERC and Fonds de la recherche en virologie et immunologie (FRVI).

post-doctoral fellowship from the Canadian Liver Foundation. J.-P.P. is an MRC scholar.

REFERENCES

1. Tassan-Dupont H, Latourneau D and Bernaudo J. P. (1996) *Nucleic Acids Res.* **24**, 1703-1705.
2. Latourneau D, M. Latourneau and Bernaudo J. P. (1997) *Nucleic Acids Res.* **25**, 1703-1705.